DEVELOPING HIGH-THROUGHPUT AMENABLE CHEMISTRY FOR CHEMICAL BIOLOGY APPLICATIONS

RAJAVEL SRINIVASAN

NATIONAL UNIVERSITY OF SINGAPORE

2009

DEVELOPING HIGH-THROUGHPUT AMENABLE CHEMISTRY FOR CHEMICAL BIOLOGY APPLICATIONS

RAJAVEL SRINIVASAN
(M.Sc., Anna University, Chennai, India)

A THESIS SUBMITTED

FOR THE DOCTOR OF PHILOSOPHY DEGREE

DEPARTMENT OF CHEMISTRY

NATIONAL UNIVERSITY OF SINGAPORE

2009

THESIS DEDICATED TO PROF. E. BALASUBRAMANIAM

(Retired Professor of chemistry, VOC College, MS University, India)

ACKNOWLDGEMENTS

First of all, I would like to express my sincere thanks to my advisor **Prof. Yao Shao Qin** for giving me an opportunity to work and learn in his laboratory. I had a wonderful chance to improve my mental discipline, leadership skills, and planning ability under his training. I always respect his straight forward approaches. It was only because of his smart ideas, deadlines and my passion towards chemistry that I was able to successfully complete most of the projects assigned to me. I am always grateful for his support and helps. I also wish for my relation/contact with Prof. Yao continues for ever.

Next, I am very grateful to **Prof. Bahulayan** (Calicut University) and **Prof. Chang Young-Tae** (NUS) for their instant and timely helps.

I am also very thankful to **Prof. V. Murugesan** (Anna University), **Dr. N. Somanathan** (CLRI), **Prof. Zhu Qing** (Zhejiang University), **Prof. Suresh** (NUS) and **Prof. S. Kumaresan** (MS University) for their kind support.

Thanks are due to my colleagues (past and present) for having provided a nice friend circle and making the working place an enjoyable one. Resmi, Souvik, Kalesh, Pengyu Hongyan, Aparna, Mahesh, Derek, Wu Hao, Ching Tian, Jiexun, Xiaohua, Mingyu, Huang Xuan, Lay Pheng, Wang Jun, Haibin, Dawn, Wang Gang, Jingyan, Su Ling and Junqi – Working with all of them have been a great experience. Special thanks to Subbu for the great chemistry discussions.

I am grateful to **Prof. Vittal** for recruiting me into NUS Chemistry graduate program and I am also thankful to NUS for awarding me the research scholarship.

I would like to express my greatest thanks to my parents **Srinivasan** and **Dhanalakshmi**, soul mate-cum-wife **Resmi**, baby **Gayathri**, baby **Srinidhi**, brothers, sisters, all **in-laws**, friends and all family members for their well wishes and ever lasting love.

TABLE OF CONTENTS

Table of Contents iii

Abbreviations vii

List of Publications xii

Summary xiv

List of Schemes xvii

List of Figures xviii

List of Tables xxii

Chapter 1 Introduction

1.1 Catalomics 1

1.2 Protein Tyrosine Phosphatases (PTPs) 2

1.2.1 Catalytic mechanism of PTPs 3

1.2.2 Inhibitor development for PTPs 4

1.3 High-throughput amenable chemistry to study PTPs 5

1.3.1 Click Chemistry 6

1.3.1.1 Click-based fragment assembly and *in situ* screening 7

1.3.1.2 Amide-bond formation and in situ screening 9

1.3.2 Solid-phase Chemistry 10

1.3.2.1 Split and mix synthesis 11

1.3.2.2 Parallel Synthesis 12

1.3.2.3 Traceless resins 13

1.3.3 Microwave Assisted Synthesis 14

1.4 Activity-based Protein profiling/fingerprinting 15

1.4.1 Components of activity-based probes 16

1.4.2 Click chemistry-based design concepts 18

1.4.3 Imaging of Enzyme activity 20

1.5 Bioimaging 21

Chapter 2 High-Throughput Assembly of Protein Tyrosine Phosphatases (PTPs) Inhibitors Using "Click Chemistry"

2.1 Summary 24

2.2 Introduction

2.2.1 Fragment-based drug discovery of PTP inhibitors 25

2.2.2 Click Chemistry as a High-throughput tool 27

2.2.3 Bidentate inhibitors against Protein Tyrosine Phosphatases 1B 28

2.2.4 PTP1B bidentate inhibitor design 29

2.3 Results and discussion

2.3.1 Chemical Synthesis of the inhibitor library 30

2.3.2 Biological screening results 33

2.4 High-throughput synthesis of 3250-member PTP inhibitor library 36

2.4.1 Traceless solid-phase synthesis of 325 azide fragments 38

2.4.2 General Procedure for the High-throughput 'click' assembly 40

2.5 Conclusion 42

Chapter 3 Solid-Phase Assembly and *in situ* screening of Protein Tyrosine Phosphatases inhibitors

3.1 Summary 44

3.2 Introduction 44

3.2.1 High-throughput amenable chemical reactions 44

3.2.2 Limitations of Wong's *in situ* screening approach 45

3.2.3 Introduction to solid-phase combinatorial library 46

3.2.4 Design of the traceless Solid-phase library 47

3.3 Results and Discussion 48

3.3.1 Chemical Synthesis of the inhibitor library 49

3.3.2 Biological screening results 51

3.4 Conclusion 54

Chapter 4 Versatile Microwave-Assisted Strategies for the Synthesis of Azide fragments

4.1 Summary 56

4.2 Introduction 56

4.2.1 Synthesis of azides – a literature review 57

4.2.2 Drawbacks of the existing methods 59

4.2.3 Design of our azide library 60

4.3 Results and Discussion 61

4.3.1 Chemical synthesis of the azide library 61

4.3.1.1 Traceless solid-phase synthesis form PS-TsCl resin 61

4.3.1.2 Solution-phase MW-assisted azidation 62

4.3.1.3 Utilization of the azides in click assembly 63

4.4 Conclusion 65

Chapter 5 Activity-based fingerprinting of Enzymes

5.1 Summary 67

5.2 Introduction 67

5.2.1 Activity-based fingerprinting 67

5.2.2 Design of our activity-based protease probes 69

5.3 Results and Discussion 71

5.3.1 Chemical synthesis of the probe 71

5.3.2 Fingerprinting experiments 73

5.4 Activity-based probe for Protein Tyrosine Phosphatases (PTPs) 75

5.4.1 Design of the probe 75

5.4.2 Chemical synthesis of the probe 77

5.5 Conclusion 79

Chapter 6 Bioimaging using small molecule probes

6.1 Summary 80

6.2 Introduction 80

6.2.1 Native chemical ligation 80

6.2.2 Small-molecular probe design 82

6.3 Results and discussions 83

6.3.1 Chemical synthesis of the probes 83

6.3.2 Imaging experiments 84

6.4 Conclusion 86

Chapter 7 Experimental Section 88

Chapter 8 Concluding Remarks 195

Chapter 9 References 196

Chapter 10 Appendix 204

ABBREVIATIONS

δ	Chemical shift in ppm
AcOH	Acetic acid
AA	Amino acid
ABPP	Activity-based protein profiling
ACC	7-Aminocoumarin-4-acetic Acid
Ala	Alanine
Arg	Arginine
Asp	Aspartic acid
Asn	Asparagine
Boc	*tert*-Butoxycarbonyl
br	Broad
tBu	*tert*-Butyl
Cy3	Cyanine dye3
Cys	Cysteine
CuAAC	Copper (I) catalyzed Azide-alkyne cycloaddition
Da	Dalton
DBU	1,8-Diazobicyclo[5.4.0]undec-7-ene
DCC	N, N'-Dicyclohexylcarbodiimide
DCM	Dichloromethane
dd	Doublet of doublet
DIC	N, N'-diisopropylcarbodiimide

DIEA	N, N'-diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DTT	Dithiothreitol
EA	Ethyl acetate
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide HCl
EDTA	Ethylenediaminetetracetic acid
ESI	Electron Spray Ionization
Gly	Glycine
Gln	Glutamine
Glu	Glutamic acid
Fmoc	9-Fluorenylmethoxycarbonyl
FMP	4-Formyl-3-methoxyphenoxy resin
HATU	O-(7-azabenzotrizol-1-yl)-1,1,3,3,tetramethyluronium hexafluorophosphate
His	Histidine
HOBT	N-Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
HT	High-throughput
Hz	Hertz
Ile	Isoleucine
Leu	Leucine

Lys	Lysine
LHMDS	Lithium bis(trimethylsilyl)amide
m	Multiplet
m-CPBA	m-Chloroperbenzoic acid
Met	Methionine
min	Minute
mm	Multiplet
mmol	Millimole
MMP	Matrix metalloproteases
MS	Mass spectrometry
MW	Microwave
NHS	N-Hydroxysuccinimide
nM	Nanomolar
NMR	N-Methylpyrrolidinone
NMP	N-methylpyrrolidone
Phe	Phenyl alanine
Pro	Proline
PTP	Protein tyrosine phosphatases
PyBOP	benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
q	Quartet
RBF	Round bottom flask
r.t.	Room temperature

s	Singlet
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
Ser	Serine
t	Triplet
TBTU	O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium tetraborofluorate
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
TLC	Thin layer chromatography
Tris	Trishydroxymethylamino methane
uv	Ultraviolet
Val	Valine
VS	Vinyl sulfone
Z	Benzyloxycarbonyl or Cbz

LIST OF PUBLICATION

1. Srinivasan; R.; Tan, L.P.; Wu, H.; Yang, P.-Y.; Kalesh, K.A,; Yao, S.Q "High-Throughput Synthesis of Azide Libraries Suitable for Direct "Click" Chemistry and in situ Screening"***, Org. Biol. Chem.***, (2009), in press**.**

2. Srinivasan, R.; Tan, L.P.; Hao, W.; Yao, S.Q "Solid-Phase Assembly and *in situ* Screening of Protein Tyrosine Phosphatases inhibitors", **Org. *Lett.,*** (2008), *10*, 2295-2298

3. Srinivasan, R.; Li, J.; Ng, S.L.; Kalesh, K.A.; Yao, S.Q. "Methods of Using Click Chemistry in the Discovery of Enzyme Inhibitors – Potential Application in Drug Discovery and Catalomics", ***Nature Protocols***, (2007), *2*, 2655-2664.

4. Srinivasan, R.; Uttamchandani, M.; Yao, S.Q. "Rapid Assembly and In Situ Screening of Bidentate Inhibitors of Protein Tyrosine Phosphatases (PTPs)" **Org. *Lett.,*** (2006), *8*, 713-716.

5. Srinivasan, R.; Huang, X.; Ng, S.L.; Yao, S.Q. "Activity-Based Fingerprinting of Proteases". ***ChemBioChem***, (2006), *7*, 32-36.

6. Srinivasan, R., Yao, S.Q.; Yeo, S.Y.D. "Chemical approaches for live cell bioimaging". ***Comb. Chem. High Throughput Screening***, (2004), *7*, 597-604.

7. Yeo, S.Y.D., Srinivasan, R.; Chen, G.Y.J.; Yao, S.Q. "Expanded utilities of the native chemical ligation reaction" ***Chem. Eur. J.***, (2004), *10*, 4664-4672.

8. Yeo, S.Y.D., Srinivasan, R., Uttamchandani, M., Chen, G.Y.J., Zhu, Q. & Yao, S.Q. "Cell-permeable small molecule probes for site-specific labeling of proteins". ***Chem. Commun.,*** (2003), 2870-2871.

9. Chattopadhaya, S.; Srinivasan; R., Yeo, S.Y.D., Chen, G.Y.J., Yao, S.Q. "Site-specific covalent labeling of proteins inside live cells using small molecule probes". ***Bioorg. Med. Chem, (***2009), 17, 981.

10. Kalesh, K.A.; Yang, P-Y.; Srinivasan, R.; Yao, S.Q. "Click Chemistry as a High-Throughput Amenable Platform in Catalomics" ***QSAR Comb. Sci.,*** (2007), *26*, 1135-1144.

11. Chattopadhaya, S.; Baker, F.B.A.; Srinivasan; R.; Yao, S.Q. "In vivo Imaging of a Bacterial Cell Division Protein Using a Protease-Assisted Small Molecule Labeling Approach" ***ChemBioChem***, (2008), *9*, 677-680.

12. Ng, S.L.; Yang, P.-Y.; Chen, K.Y.-T.; Srinivasan, R.; Yao, S.Q. "Click" Synthesis of Small Molecule Inhibitors Targeting Caspases. ***Org. Biomol. Chem.***, (2008), 2008, 6, 844-847.

13. Yang, P-Y.; Wu, H.; Lee, M.Y.; Xu, A.; Srinivasan, R.; Yao, S.Q. "Solid-Phase Synthesis of Azidomethylene Inhibitors Targeting Cysteine Proteases". ***Org. Lett.***, (2008), 10, 1881–1884.

POSTERS PRESENTED AT CONFERENCES

1. Srinivasan, R.; Yang, P-Y.; Yao, S.Q. "Highthroughput amenable Chemistry in Catalomics" A*STAR-Noyori Forum Joint Symposium on Organic Chemistry. Singapore, May 2007.

2. Srinivasan, R.; Uttamchandani, M.; Yao, S.Q. *'Click'-assembly of bidendate PTP1B inhibitors:* Presented at the Singapore International chemical conference (SICC-4). Singapore, December 2005.

3. Srinivasan, R.; Huang, X.; Yao, S.Q. *Chemical Biology of Phosphatases:* Presented at the Faculty Graduate Congress, National university of Singapore. Singapore, Sep 2005.

4. Srinivasan, R.; Huang, X.; Yao, S.Q. *Activity based fingerprinting of Proteases:* Presented at the First Singapore Mini-Symposium on Medicinal Chemistry: Advances in Synthesis and Screening. Singapore, July 2005.

5. Srinivasan, R.; Yao, S.Q. *Design and Synthesis of cell permeable Small-molecule probes for live-cell imaging:* Presented at the Singapore International chemical conference (SICC-3). Singapore, December 2003.

ORAL PRESENTATION

1. Srinivasan, R.; "Click" Chemistry in Catalomics – Faculty of Science Graduate Congress, NUS, Singapore, Sep - 2007.

2. Srinivasan, R.; "Highthroughput amenable Chemistry in Catalomics" A*STAR-Noyori Forum Joint Symposium on Organic Chemistry. Singapore, May 2007.

3. Srinivasan, R.; "Highthroughput amenable Chemistry in Catalomics", 3rd MPSGC, Kuala lampur, Malaysia, Dec – 2007

AWARDS

1. ***Best Graduate Researcher*** in Chemistry, NUS – Sep 2007

2. ***Top Graduate Researcher*** in the Faculty of Science, NUS – Sep 2007

SUMMARY

Proteins are very crucial for cellular and metabolic functions. They play crucial roles like cell division, cell migration, cell-cell communication, signal transduction, cell death etc., and malfunction of proteins leads to major human diseases like AIDS, Cancer, Alzheimer's, diabetes etc. Enzymes account for more than 20% of the drug targets and from an analysis of the FDA orange book, it has been estimated that atleast 370 marketed drugs work by inhibiting an enzyme. There are around 18-29% of the eukaryotic genomes which encode enzymes. However, little is still understood about the physiological role, substrate specificity and downstream targets of enzymes, which necessitates their study in a high-throughput manner. "Catalomics" is the emerging sub-field of Chemical Biology in which one aims at studying enzymes at the organism wide scale by employing high-throughput chemistry and technology. High-throughput platform offers an efficient and rapid analysis of proteome on a global scale. The major part of my Ph.D research concentrates on the development or fine-tuning of high-throughput amenable chemical platform to study enzymes especially Protein Tyrosine Phosphatases (PTPs), the last part of my research work is focused on Activity-based profiling and Bioimaging.

One of the main challenges in the field of Catalomics is the development of high-throughput (HT) amenable chemical reactions that allow rapid synthesis of diverse chemical libraries for the interrogation of different classes of enzymes. High-throughput (HT) amenable reactions are mainly characterized by near-perfect, modular and robust and biocompatible nature. Among the HT amenable chemistry tools, we have explored on Click chemistry, Solid-phase chemistry and Microwave (MW) assisted reactions

synthesis for enzyme inhibitor libraries and fragments. We have successfully adopted the click chemistry platform to construct a 66-member isoxazole based bidentate inhibitor library targeting against various PTPs. Subsequent *in situ* screening revealed a potential inhibitor against PTP1B (IC_{50} = 4.7 μM) which is comparable to that of a most potent cell active PTP1B inhibitor known. Following the results, next we successfully designed and developed a traceless solid-phase methodology to construct different azide fragments, the azides were of high-quality and 100s of azides can be easily made within a week. These azide library was used constructed a library of 3250-member isoxazole-based bidentate inhibitor library targeting various PTPs in a high-throughput fashion.

Next, we successfully adopted out traceless solid-phase methodology to rapidly assemble PTP inhibitors using amide bond-formation reaction. The highlight of our method is that again no purification was required after the synthesis and cleavage from the solid-support. The inhibitors were of high-purity and were suitable for *in situ* screening. By screening these inhibitors against PTP1B, we have uncovered a candidate molecule which possesses an inhibition of K_i = 7.0 μM against PTP1B.

We have devised another simple and practical microwave-assisted strategy for the conversion of readily available alcohol building blocks into azides. In the first route tosyl resin was used for the first time to synthesize azides in a catch and release approach. In the second approach, alcohol building blocks were readily converted into azides via MW-assisted azidation of tosylates/mesylates/chlorides. After a simple purification procedure, the azides synthesized from the above methods were suitable for directly click chemistry applications.

The next part of the thesis work was focused on the synthesis of 16 different Activity-based probes; these probes were used to generate unique substrate fingerprint profiles of proteases in gel-based proteomic experiments. We have also synthesized a tri-functional PTP probes. In the last part of the thesis work, I have presented the design and synthesis of a set of cell-permeable small molecule probes which were utilized to modify (label) selectively the N-terminal cysteine proteins (*in vivo*) in bacterial cells by the well-known native chemical ligation.

LIST OF SCHEMES

Scheme 2.1 Synthesis of the alkyne building blocks 31

Scheme 2.2 Synthesis of Azide building blocks 32

Scheme 2.3 Click assembly of PTP inhibitors 33

Scheme 2.4 Traceless solid-phase synthesis of azide libraries 39

Scheme 2.5 Synthesis of the alkyne building blocks (set II) 40

Scheme 3.1 Traceless solid phase synthesis of PTP1B inhibitors 51

Scheme 4.1 (a) Solid-phase synthesis of azides from P-TSCl resin (b) Synthesis of azides via tosylates (c) Synthesis of azides via mesylates/chlorides. 63

Scheme 4.2 Utilization of the azides in click assembly of PTP inhibitors 65

Scheme 5.1 Synthesis of the activity-based protease probes 71

Scheme 5.2 Synthetic scheme for PTP tri-functional activity-based probes 78

Scheme 6.1 Synthesis of the cell-permeable probes for live cell labeling 84

LIST OF FIGURES

Fig. 1.1 Biological role of Protein Tyrosine Phosphatases (PTPs) 2

Fig. 1.2 Classification of PTP super family 3

Fig. 1.3 Catalytic mechanism of PTPs 4

Fig. 1.4 Various Pharmacophores targeting PTPs 5

Fig. 1.5 High-throughput Chemistry tools 6

Fig. 1.6 Cu(I) catalyzed azide-alkyne ligation 7

Fig. 1.7 Fragment-based assembly and in situ screening 8

Fig. 1.8 'Click' inhibitors targeting various enzyme classes 9

Fig. 1.9 (a) Amide-bond formation reaction (b) Inhibitors assembled by amide-bond formation reaction 10

Fig. 1.10 General principle of solid-phase synthesis 11

Fig. 1.11 Schematic representation of split and mix synthesis 12

Fig. 1.12 Schematic representation of parallel synthesis 12

Fig. 1.13 Application of Aldehyde traceless resin 13

Fig. 1.14 (a) DHPM cores synthesized by high-speed automated MW technology (b) Microwave assisted solid phase synthesis of bicyclic dihydropyrimidones 15

Fig. 1.15 Representation of Activity-based labeling strategy 16

Fig. 1.16 Components of Activity-based probes 17

Fig 1.17 **(a)** 'Click' based two-step labeling approach **(b)** MJE3, natural product inspired activity-based probe **(c)** Labeling of cytochrome P450 **(d)** Labeling of proteases metalloproteases 19

Fig. 1.18 Activity-dependent labeling of cysteine protease by qABP 20

Fig. 1.19 **(a)** Site-specific labeling of tetracysteine motif by FlAsH **(b)** NTA probe coordinating to Ni (II) ion. **(c)** Labeling of hAGT fusion protein with BG derivative 23

Fig. 2.1 Different fragment-based approaches to discovery PTP inhibitors **(a)** SAR by NMR method **(b)** 'Tethering' approach **(c)** Substrate activity screening method 26

Fig 2.2 Click chemistry and *in situ* screening of inhibitor discovery 27

Fig. 2.3 Bidentate inhibitors against PTP1B 28

Fig. 2.4 **(a)** Click assembly of PTP1B inhibitors **(b)** Mode of binding of the PTP1B bidentate inhibitor 29

Fig. 2.5 IC_{50} graphs for selected inhibitors against PTP1B and TCP TP 34

Fig. 2.6 Structures of the potent inhibitors identified from the screening 35

Fig. 2.7 Schematic representation of high-throughput azide synthesis 37

Fig. 2.8 Schematic representation of 3250-member 'click' assembly 40

Fig. 3.1 **(a)** Wong's solution-phase and *in situ* screening approach **(b)** Stable benzotriazole by-products in amide-bond formation reaction. 46

Fig. 3.2 Solid-phase synthesis of Indinavir® analogues 47

Fig. 3.3 Yao's solid-phase and *in situ* screening approach 47

Fig 3.4 Graphical representation of our solid-phase strategy 48

Fig. 3.5 **(a)** Inhibitor design- solid-phase amide bond formation reaction **(b)** Most potent inhibitor for the compound library 49

Fig. 3.5 Six candidate hits identified against PTP1B. 52

Fig. 3.6 Dixon plots for determination of K*i* values of three representative inhibitors against PTP1B 53

Fig 4.1 **(a)** Solid-phase synthesis of aryl azides from aryl triazenes **(b)** Solid-phase synthesis of heteroaryl azide from heteroaryl sulfone 58

Fig 4.2 Various methods of synthesizing azides from commercially available building blocks 59

Fig. 4.3 MW-assisted synthesize of azides followed by direct click chemistry. **(a)** Solid- phase catch and release approach. **(b)** Solution-phase approach 61

Fig. 5.1 Structure of the AB-protease probes developed by Yao's and Cravatt's lab 69

Fig. 5.2 Structure of the activity-based protease probe 70

Fig. 5.3 Mechanism of labeling of the protease probe 71

Fig. 5.4 **(a)** Affinity-Based labeling of Trypsin with probes 16 probes in the descending order. **(b)** Fingerprints of various proteases with probes **5-1** to **5-16**. 75

Fig. 5.5 Structure of the PTP tri-functional probe 76

Fig. 6.1 Native chemical ligation 82

Fig. 6.2 Chemoselective reaction between a thioester-containing probe and an N-terminal Cys protein in a living cell 83

Fig 6.4 Site-specific labeling of N-terminal cysteine proteins with the small molecule Probes 86

Fig. 7.1 Structure of the fluorogenic substrate used in the assay 99

Fig. 7.2 Inhibition profiles of the 66-member library against 6 phosphatases 101

Fig. 7.3 IC_{50} graphs for various 'click' inhibitors against different enzymes. 103

Fig. 7.4 List of amine building blocks 103

Fig. 7.5 List of acid building blocks. 115

Fig. 7.6 List of sulfonyl chloride building blocks 120

Fig. 7.7 List of 35 different amines used for the reductive amination 138

Fig. 7.8 Graphs for determining IC_{50} values of selected library members against PTP1B 145

Fig. 7.9 List of alcohol blocks used in route A 147

Fig. 7.10 List of azides synthesized from PS-TsCl resin 148

Fig. 7.11 List of alcohol blocks used in tosylation 150

Fig. 7.12 List of tosylates synthesized 151

Fig. 7.13 List of azides synthesized via tosylates 153

Fig. 7.14 List of alcohols used for mesylation 155

Fig. 7.15 List of mesylates/chlorides synthesized 166

Fig. 7.16 List of azides synthesized via mesylation 159

Fig. 7.17 Labeling of various concentrations of trypsin 177

Fig. 7.18 Labeling of Trypsin with various probe concentrations 177

Fig. 7.19 SDS-PAGE of 13 different proteases labeled with **Probes 5-1** to **5-16** 178

Fig. 7.20 SDS-PAGE of purified N-terminal cysteine EGFP labeled 193

List of Tables

Table 2.1 IC_{50} (in μM) of 6 selected inhibitors. 36

Table 2.2 Optimization conditions of the Click reaction 41

Table 3.1 Inhibition of the six selected inhibitors 54

Table 4.1 List of all azides and yields synthesized by three different approaches 64

Table 7.1 Reaction conditions for click reaction 128

Table 7.2 Summary of ESI-MS results of the probes **5-1** to **5-16** 176

Table 10.1 Summary of traceless azide characterization 213

Table 10.2 Summary of the triazole characterization 221

Table 10.3 Summary of characterization and IC_{50} of traceless PTP1B inhibitors 223

Chapter 1

Introduction

1.1 Catalomics

After the successful completion of the major human genome project[1], more attention has turned towards proteomic research rather than genomic research. Proteins are very crucial for cellular and metabolic functions. They play crucial roles like cell division, cell migration, cell-cell communication, signal transduction, cell death etc and malfunction proteins leads to major human diseases like AIDS, cancer, Alzheimer's, diabetes etc. It has been estimated that enzyme account for more than 20% of the drug targets and from an analysis of the FDA orange book, it has been estimated that at least 370 marketed drugs work by inhibiting an enzyme[2]. There are around 18-29% of the eukaryotic genomes which encode enzymes. However, little is still understood about the physiological role, substrate specificity and downstream targets of enzymes, which necessitates their study in a high-throughput manner. "Catalomics" is the emerging sub-field of Chemical Biology in which one aims at studying enzymes at the organism wide scale by employing high-throughput chemistry and technology[3]. High-throughput platform offers an efficient and rapid analysis of proteome on a global scale. My Ph.D research mainly concentrates on the development or fine-tuning of high-throughput amenable chemical platform to study enzymes especially Protein Tyrosine Phosphatases (PTPs).

1.2 Protein Tyrosine Phosphatases (PTPs)

PTPs are a class of signaling enzymes which catalyze the dephosphorylation of the phosphotyrosine residues in their protein substrate[4] (**Fig. 1.1**). Protein phosphorylation and dephosphorylation reactions are employed by living organisms for the regulation of innumerable cellular processes. (e.g., Signal transduction, a process in which cell converts one form of signal into another). Phosphorylation states are governed by Protein Kinases whilst dephosphorylation states are governed by Protein Phosphatases. Malfunction of PTPs leads to various human diseases like Cancer, diabetes, obesity etc[5] and thus PTP is an important drug target. More than 100 identified PTPs are encoded in genome and at least 400 remains to be identified.

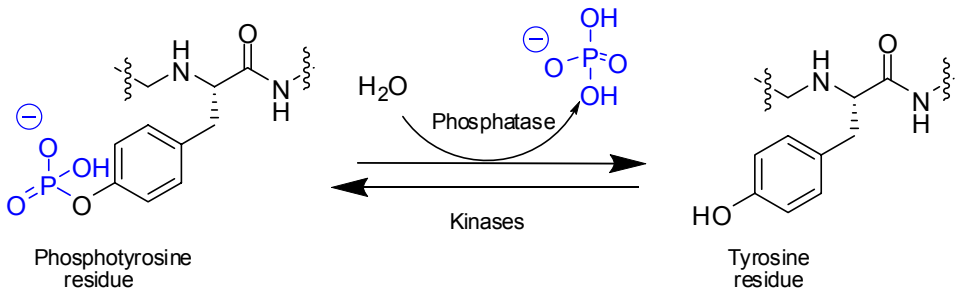


Fig. 1.1 Biological role of Protein Tyrosine Phosphatases (PTPs)

The PTP super family can be classified into 3 main sub-families according to the amino acid sequence of PTP catalytic domain[5] (**Fig. 1.2**).

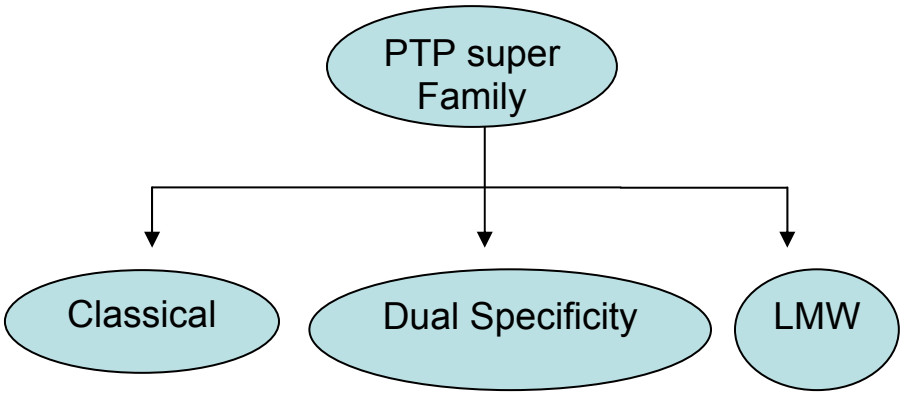


Fig. 1.2 Classification of PTP super family

(a) Classical Phosphatases:

Classical phosphatases hydrolyse the pTyr residues in the protein substrates and they are further classified as intracellular and receptor like phosphatases. e.g. PTP1B

(b) Dual specificity phosphatases:

They dephosphorylate the pTyr as well as pSer/pThr residues in their protein substrates. Their active site motif is given by the sequence (H/V)C(X)5R(S/T).

e.g. VHR, Cdc subfamily, PTEN (It can dephosphorylate phosphoinositides as well)

(c) Low Molecular Weight Phosphatases (LMW):

They are the family of 18 kDa proteins which are involved in the dephosphorylation of the growth factor receptors. e.g., LMW-DSP4

1.2.1 Catalytic mechanism of PTPs

All the PTPs share the common catalytic mechanism. The active site sequence is **C**(X)5**R**(S/T) which is also called the PTP signature motif[6]. Cysteine (C) and Arginine (R) are invariant and necessarily essential for catalysis. The Arg (221) in the active site of the PTP stabilizes the negatively charged phosphate group of the substrate. Upon

substrate binding the enzyme undergoes a conformation change which further brings the substrate close, followed by the attack of the Cys(215) on the phosphate group, a water molecule acts as a nucleophile in the breaking of the S-P bond resulting in the dephosphorylation of the substrate (**Fig. 1.3**).

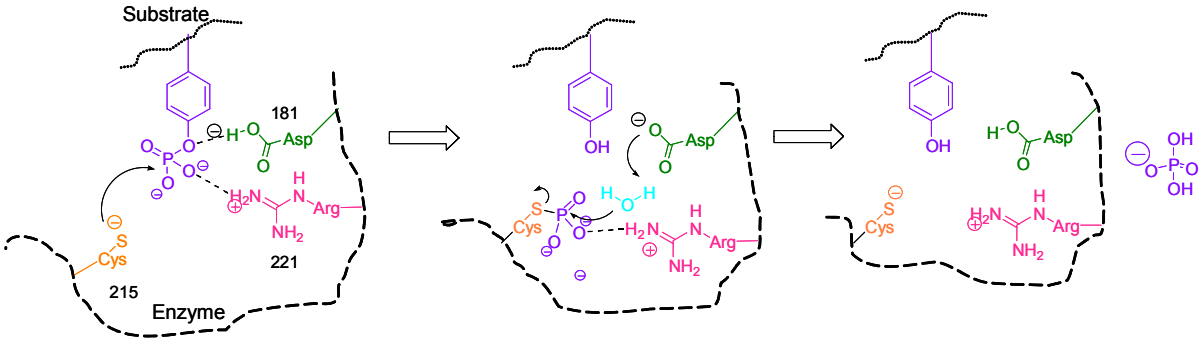


Fig. 1.3 Catalytic mechanism of PTPs

1.2.2 Inhibitor development for PTPs

Various drug companies as well as academic groups have been working on the development of various PTP inhibitors based on the pTyr mimetic. The most common ones are listed (**Fig. 1.4**). However, due to the negative charge(s) on the molecule most of the inhibitors suffer from poor bioavailability and poor cell permeability. Also, there lies a major challenge in the development of an inhibitor specific towards a particular PTPs, which is discussed in the following chapter. Recent generation of PTP inhibitors[5] like the oxamic acid based as well as the isoxazole based inhibitors are proven to be cell permeable and bioavailable. We have adopted and developed some high-throughput amenable chemistry tools to develop better inhibitors against PTPs, which is discussed in the following section.

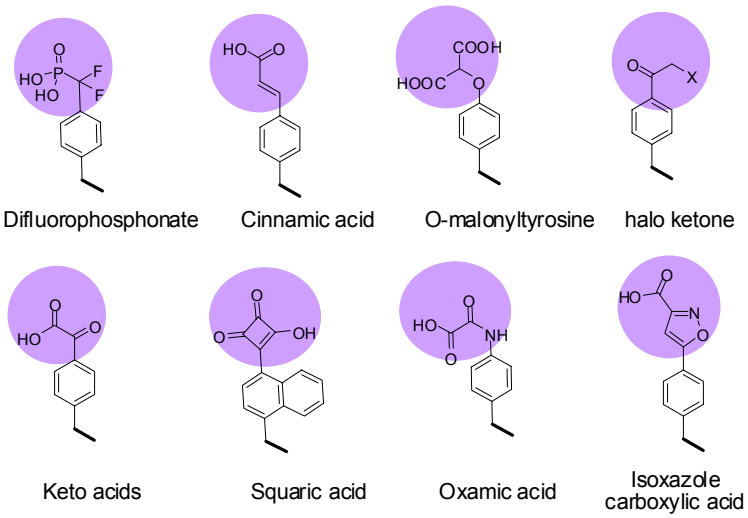


Fig. 1.4 Various Pharmacophores targeting PTPs.

1.3 High-throughput amenable chemistry to study PTPs

One of the main challenges in the field of Catalomics is the development of high-throughput (HT) amenable chemical reactions that allow rapid synthesis of diverse chemical libraries for the interrogation of different classes of enzymes. High-throughput amenable reactions are mainly characterized by near-perfect, modular and robust and biocompatible nature. Among the HT amenable chemistry tools (**Fig. 1.5**), we have explored on Click chemistry, Solid-phase chemistry and Microwave (MW) assisted reactions for synthesis for the enzyme inhibitor libraries and fragments.

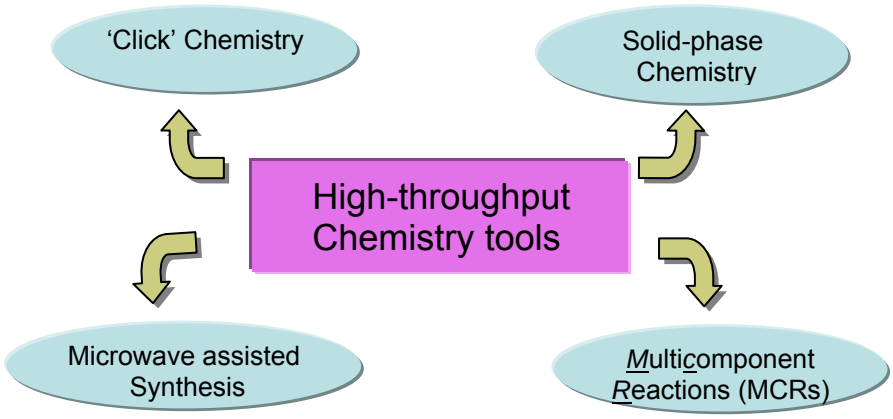


Fig. 1.5 High-throughput Chemistry tools

1.3.1 Click Chemistry

Sharpless et al have defined Click Chemistry as a set of powerful, virtually 100% reliable, selective reactions for the rapid synthesis of new compounds *via* heteroatom links (C-X-C)[7]. Among the various click reactions the reaction between a terminal alkyne and an azide in the presence of Cu(I) catalyst generate exclusively the 1,4-disubstituted 1,2,3-triazoles discovered independently by Sharpless[8] and Meldal[9] in the most celebrated one and more generally referred to as Click chemistry (**Fig. 1.6**). The reaction is characterized by its high chemoselectivity, modularity, near-perfect yield and biocompatibility in the aqueous conditions are typically employed in the reaction, which renders the products 'ready-to-use' without further purifications. As a result, click chemistry has become an attractive tool in many research fields ranging from materials sciences, biology to medicinal chemistry/chemical biology. It is for the same reasons that click chemistry has emerged as an integral part of the drug discovery pipeline by providing a high-throughput amenable chemical reaction platform for compound synthesis.

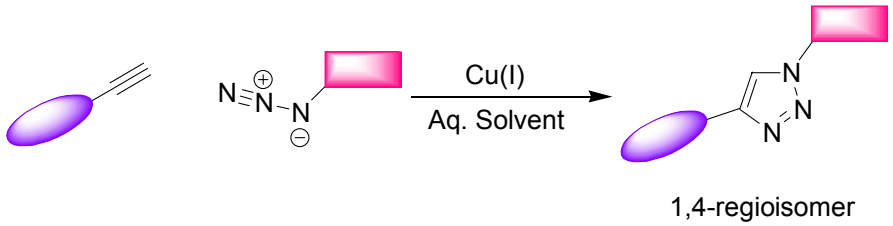


Fig. 1.6 Cu(I) catalyzed azide-alkyne ligation

Other salient features of Click Reaction[10]

1. Highly exothermic reaction (ΔH = -45 to -55 Kcal/mol)
2. High kinetic barrier – due to the high energy content of the azide and alkyne (25 kcal/mol)
3. In the Presence of Cu(I) the reaction is not significantly affected by the steric and electronic properties of the groups attached to the azide and alkyne reactive centers.
4. The reaction is unaffected by water and other functional group, the rate of Cu(I) catalyzed is 10^7 times faster than the uncatalyzed one
5. Triazole ring is very stable. Inert to oxidation, hydrolysis, and reduction conditions even at high temperature.
6. Strong Dipole moment, aromatic, good H-bonding acceptor.

1.3.1.1 Click-based fragment assembly and *in situ* screening

Fragment-based assembly is a recently developed drug discovery approach which enables high-throughput identification of small molecule inhibitors using a minimal number of compounds as building blocks[11]. The approach is powerful especially against protein targets which possess multiple binding pockets in their active sites (e.g.,

Proteases, Phosphatases, and Kinases etc). Currently, there are mainly three methods used to assist the assembly of compounds: (1) the NMR-based SAR strategy[12]; (2) the tethering method developed by Wells et al [13] and (3) the in situ screening method developed by Wong and co-workers[14] which was largely based on the "click chemistry" pioneered by Sharpless et al. **(Fig. 1.7)**. Among them, the click chemistry approach is highly versatile in that it requires neither specialized equipment nor mutations in the target proteins.

Properties of in situ screening

1. Reactions are usually carried out at microscale level
2. Water compatible or non-toxic solvent should be used for the reaction
3. No protection group required
4. Products are not isolated or purified, Ready-to-use products.
5. Only high yielding reactions are suitable.

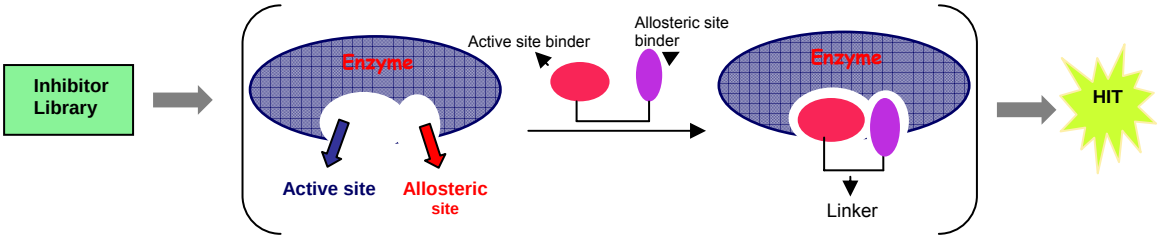


Fig. 1.7 Fragment-based assembly and in situ screening.

Till to date various groups including ours have utilized this powerful tool for the successful discovery of inhibitors against PTP1B[15,16], Matrix metalloproteases (MMPs)[17], HIV protease[18], SARS 3CL protease[19], sulfotransferase[20], α-1,3-fucosyltransferase[21], acetylcholinesterase[22] and caspases[23] (**Fig. 1.8**).

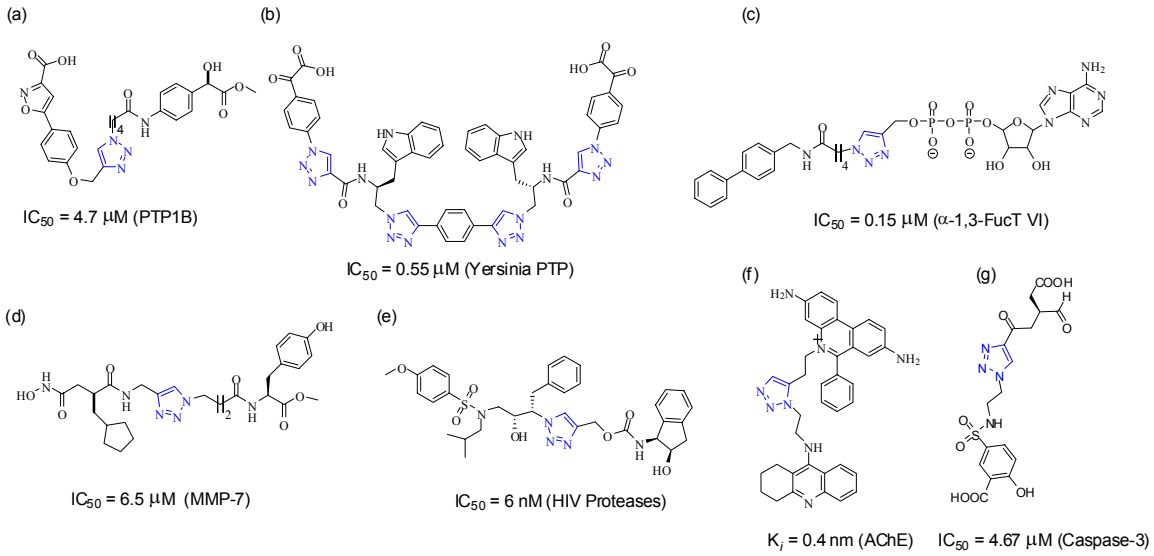


Fig. 1.8 'Click' inhibitors targeting various enzyme classes.

1.3.1.2 Amide-bond formation and *in situ* screening

The reaction between an acid and an amine to generate amide using suitable activating/coupling reagent is also a near-perfect reaction and can be amenable to high-throughput synthesis and in situ screening[14] (**Fig. 1.9a**). The most common coupling reagents used is EDC, HATU, HBTU, Bybop etc. The amide-forming reaction between an amine and an acid is one of the most efficient reactions known, and has recently been applied successfully in the rapid discovery of inhibitors against a number of enzymes, such as cysteine proteases[24], HIV proteases[25], HIV-1 dimerization[26] inhibitor, β-aryl sulfotransferase[27], α-fucosidases[28] and SARS-3CL proteae[29] (**Fig. 1.9b**). This chemistry has some key characteristics of a high-throughput amenable reaction, in that quantitative product formation can be achieved using powerful acylating/coupling reagents. As such, in situ biological screening may be carried out directly without product purification. But the method is severely limited by the presence of coupling reagents and, more often than

not, byproduct and excessive starting materials in the reaction, a subsequently during in situ screening may lead to false positive results.

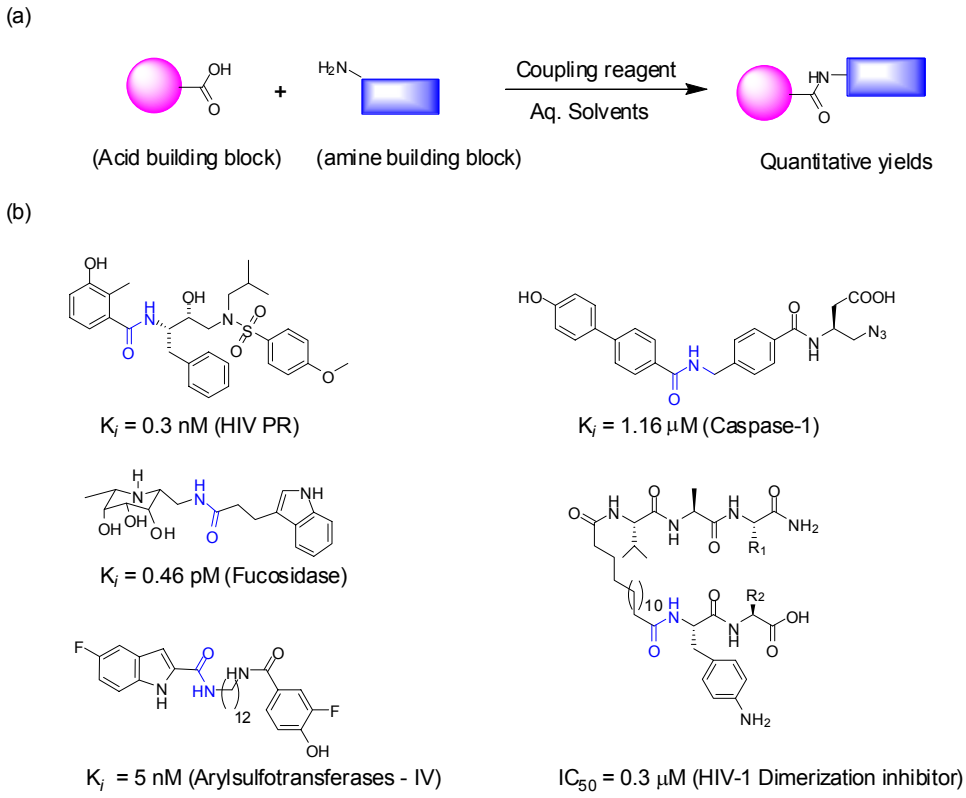


Fig. 1.9 (a) Amide-bond formation reaction (b) Inhibitors assembled by amide-bond formation reaction

1.3.2 Solid-phase synthesis

The next high-throughput tool that we have explored in Catalomics study is the solid-phase library synthesis (**Fig. 1.10**). Most of the combinatorial libraries are generated by solid-phase method. In 1963 Merrifield published the first landmark paper on solid-phase synthesis of a tetrapeptide[30] and Ellman was the first to report the solid-phase combinatorial synthesis of non-peptide based small molecules, benzodiazepines[31]. Solid phase synthesis have several advantages over solution phase synthesis, which are listed below

1. Can be automated easily and amenable to high-throughput synthesis
2. Isolation of the compound is very simple, usually by simple filtration
3. High product purity
4. Easier to generate a large library of compound in a shorter time
5. In some cases the reaction can be forced to completion by using excess of reagent
6. In some cases (especially peptide synthesis), solubility problems are minimized

However, the main limitations of solid-phase synthesis are difficulty in monitoring the reaction, usage of excess reagents and solvents, difficult in multistep stereo-selective synthesis.

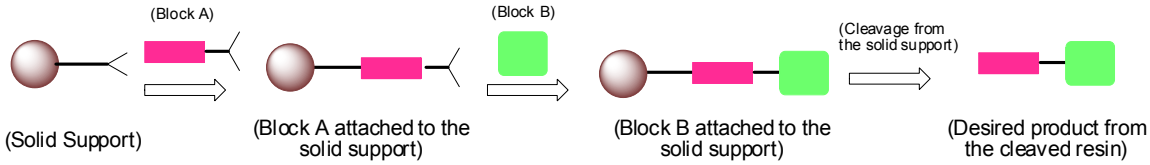


Fig. 1.10 General principle of solid-phase synthesis

Two general methods include in combinatorial synthesis is the split and mix synthesis and parallel synthesis.

1.3.2.1 Split and mix synthesis

The split and mix method consists of three processes: (a) splitting, (b) coupling (or other reaction) and (c) mixing[32] (**Fig. 1.11**). First, the resin beads are split into multiple reaction vessels and carried out the reaction with different individual compounds. The polymer beads are randomly mixed after the reaction. Thus splitting and mixing are done alternatively. The advantage of this method is easier and faster when compared to the parallel synthesis method and the compounds in the library grow exponentially with the number of steps.

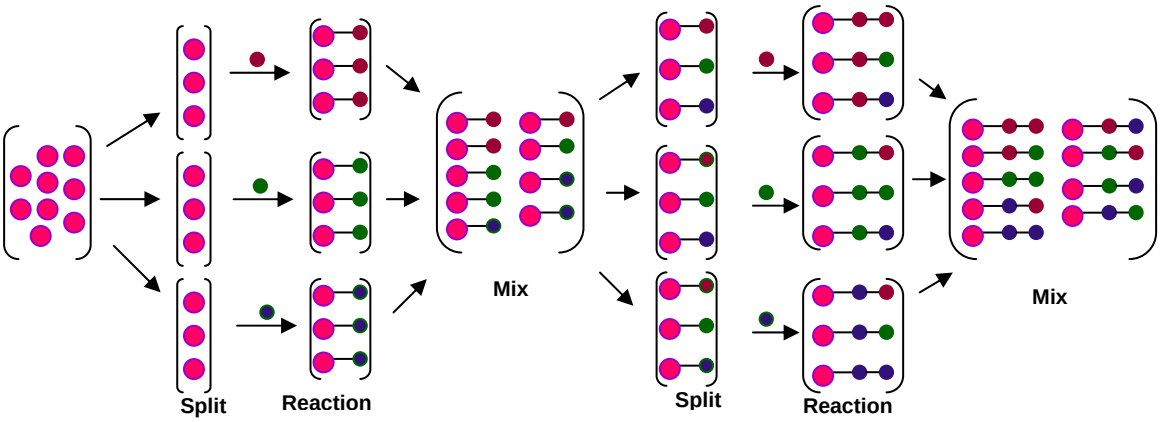


Fig. 1.11 Schematic representation of split and mix synthesis

1.3.2.2 Parallel synthesis method

It is a very simple approach in which the desired compounds are synthesized in individual reaction vessels and therefore all products are pure, separated and well defined[33] (**Fig. 1.12**). However the main bottle neck of this method is the library size is usually very small. The synthesis can be either manual or automated.

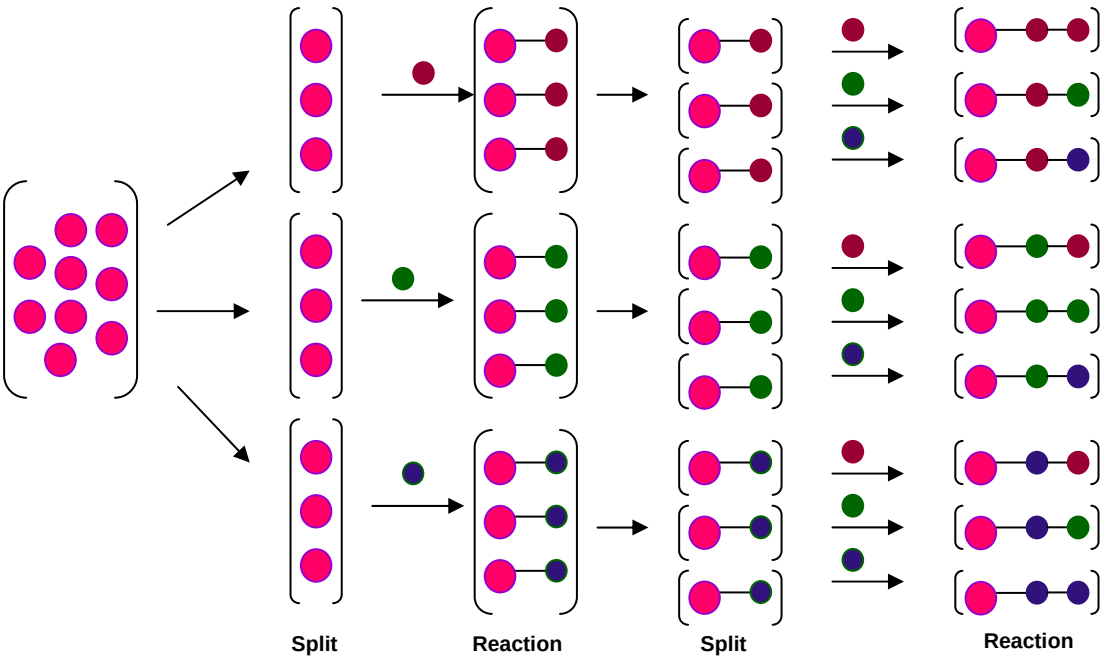


Fig. 1.12 Schematic representation of parallel synthesis

1.3.2.3 Traceless resins

We have developed a traceless solid phase methodology for the high-throughput synthesis of PTP inhibitors (**Fig. 1.13**). Traceless synthesis can be defined as a synthetic route which yields compounds composed of only from atoms inherent to the particular target molecule[34]. Synthesis using solid-phase traceless resins leaves no residues on the final compound after the cleavage from the solid support. We have explored the solid phase amide bond formation reaction and urea formation reaction using aldehyde traceless resin (PL-FMP)[35] and for the first time we have synthesized a library of azides using the traceless tosyl resin. The former traceless method can be adapted to the synthesis of various drug like fragments and other important organic intermediates like urea, guanidines etc.

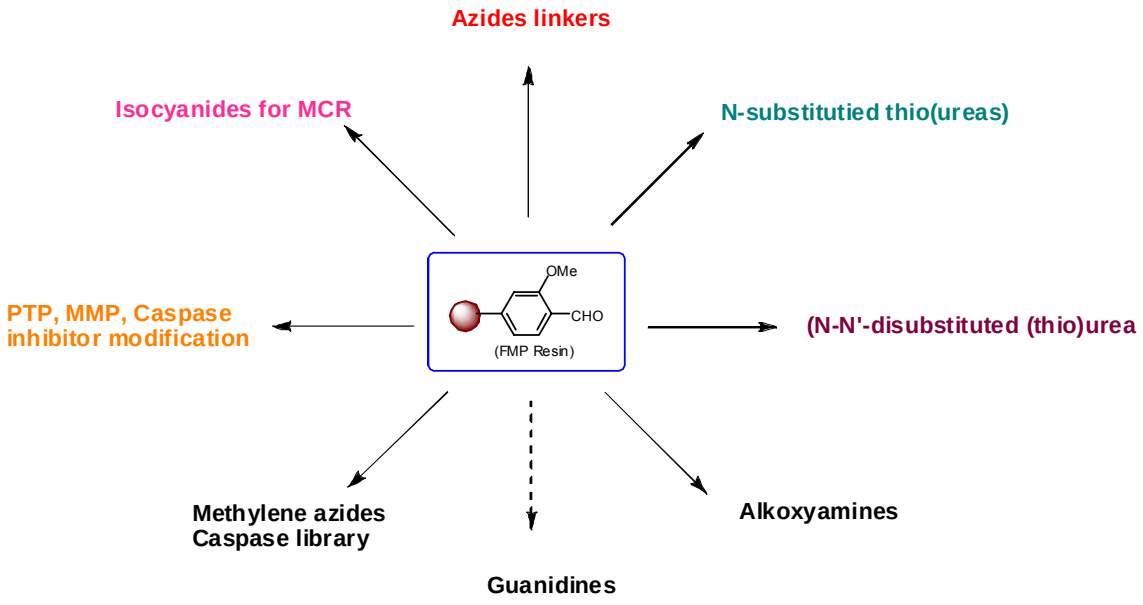


Fig. 1.13 Application of Aldehyde traceless resin

1.3.3 Microwave-Assisted Synthesis

The next high-throughput chemistry tool that we have explored for rapid generation of compound libraries is the microwave assisted organic synthesis (MAOS). Traditional method of heating reactions using oil bath, mantles etc is inefficient because it depends on the thermal conductivity of the various materials that must be penetrated and the reaction of the reaction vessels is higher than the reaction medium whilst microwave heating is internal and depends on the dipole moment of the molecules[36]. Microwave assisted organic reactions have been proven to an invaluable tool for high-throughput chemists because of the following advantages

1. Dramatic reduction in reaction time, the reaction time can be reduced from days to minutes
2. The overall process is energy efficient than the oil-bath method
3. Precise monitoring and control of temperature and pressure possible
4. Can be automated
5. The choice of the solvent is not governed by the boiling point
6. Most of the reactions are cleaner and provide higher yields.
7. Higher reaction temperature can by obtained (sealed tubes should be used).
8. Facilitate the discovery of novel reaction pathways

Recently Kappe et al have generated a library of Biginelli products using automated sequential microwave assisted chemistry[37]. A total of 48 different dihydropyrimidine (DHPM) analogues were synthesized by automated addition of building blocks and subsequent sequential microwave irradiation of each process vial. The desired products ware obtained in an average yield of 52 % and >90 purity (**Fig. 1.14a**).

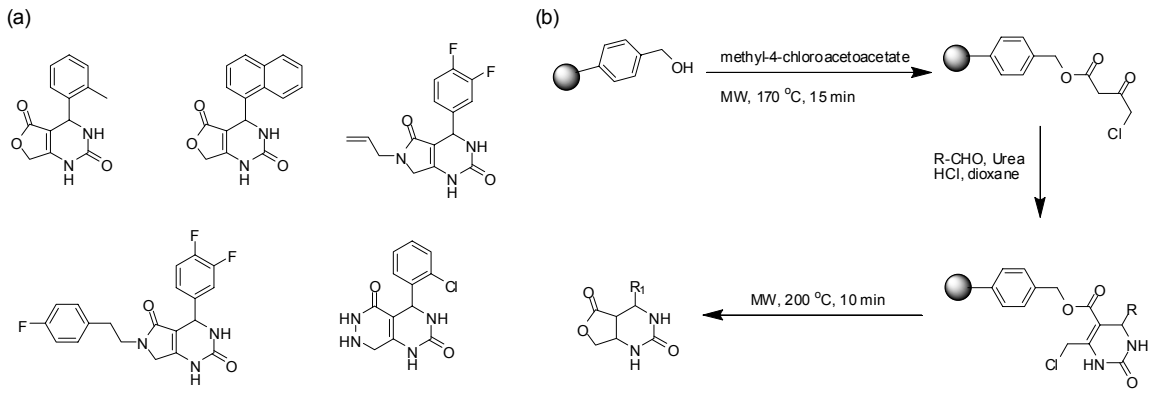


Fig. 1.14 (a) DHPM cores synthesized by high-speed automated MW technology **(b)** Microwave assisted solid phase synthesis of bicyclic dihydropyrimidones.

In another recent report, Kappe et al have synthesized a library of bicyclic dihydropyrimidones by traceless solid-phase approach using MW-assisted cyclization and cleavage[38] (**Fig. 1.14b**). They were able to achieve 99 % conversion within 15 min at 170 °C, as confirmed by on-bead FTIR analysis.

1.4 Activity-based Protein profiling/fingerprinting

In order to accelerate the functional analysis of a cell's complete protein repertoire, high-throughput techniques capable of genome-wide studies of proteins are essential. Recently, the activity-based profiling of proteins has proven to be a powerful tool in proteomic studies, whereby subclasses of enzymatic proteins could be selectively identified[39]. This strategy takes advantage of mechanism-based probes that react with different classes of enzymes, leading to the formation of covalent probe–protein complexes which are readily distinguished non-reactive proteins in a crude proteome mixture from other non-reactive proteins in a crude proteome mixture. **Fig. 1.15** shows a general approach to activity-based enzyme profiling using a fluorescently activity-based

probe. The probe selectively labels a particular class of active enzyme present in a complex proteome with many other different proteins. The labeled enzyme can be separated by SDS-PAGE and visualized by fluorescent imaging, which could be further characterized by mass spectrometry.

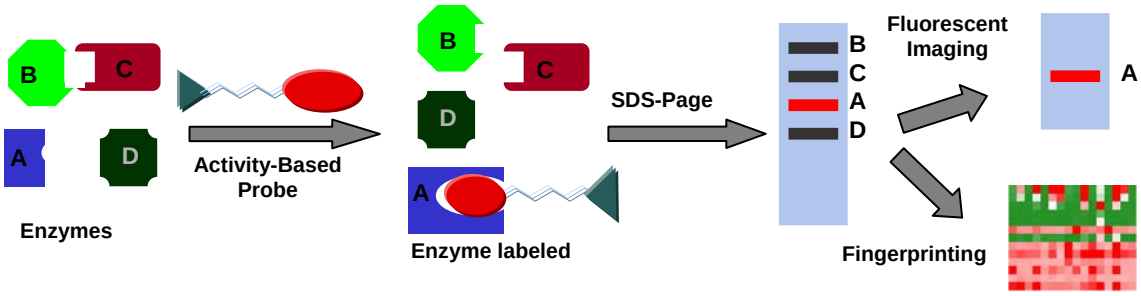


Fig. 1.15 Representation of Activity-based labeling strategy

Highly specific probes offer invaluable insights into the topology of the enzyme active site and catalytic mechanism of particular enzyme sub-class; whilst broad-based probes can label most proteins having the same activity and have more application in profiling enzymes on a global scale.

1.4.1 Components of activity-based probes

An activity-based probe is generally made of three units (**Fig. 1.16**)

1. Reactive unit (e.g. Fluoromethyl ketone[40], vinyl sulfone[41], epoxides[42], fluorophosphonate[43], α-bromophenyl phospnotate[44])
2. Linker (e.g. peptide fragment, polyethene linker or alkyl linker)
3. Reporter unit (e.g. biotin, fluorescent dye like Cy3, rhodamine etc)

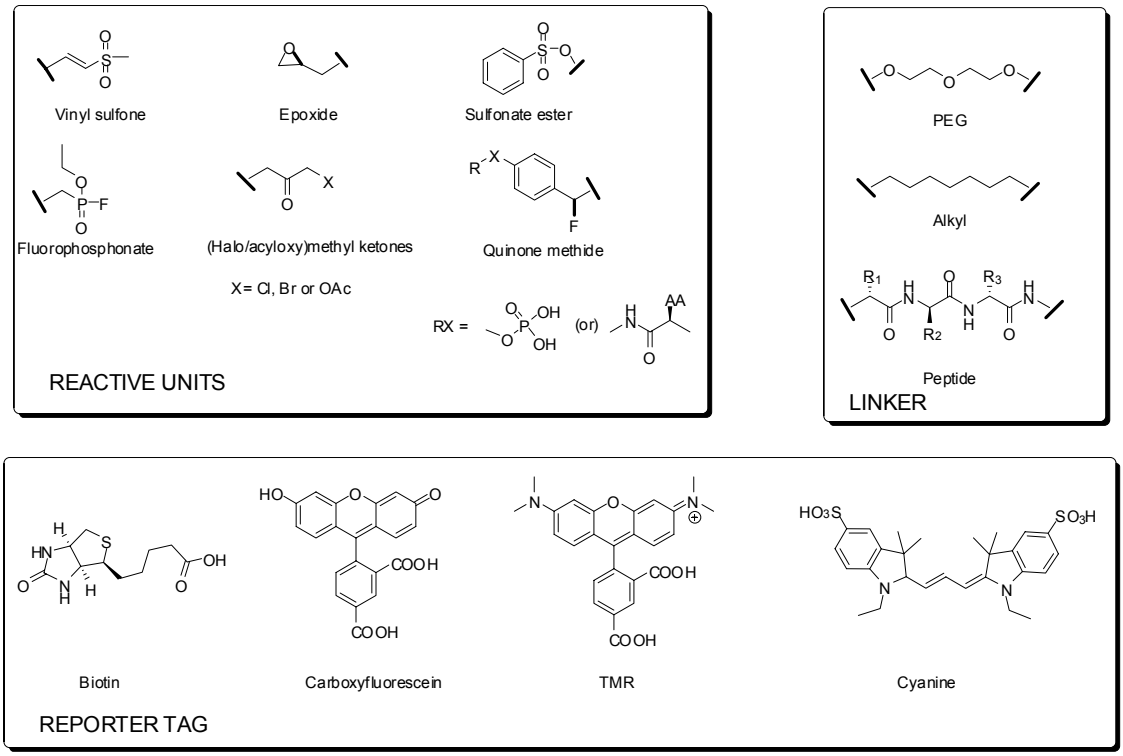


Fig. 1.16 Components of Activity-based probes

The reactive unit can be a suicide/mechanism-based inhibitor of one particular class of enzyme which serves as a warhead to covalently modify the enzyme. The reactive units are generally electrophilic chemical groups. The linker forms a bridge between the reactive unit and the reporter unit, the length of the linker is very important to keep the reporter unit away for the active site. The reporter unit can be a fluorescent or radio or biotin tag, the tag should be compatible with standard SDS-PAGE techniques. Biotin facilitates detection by simple western blot approaches. Biotin label can also be used to enrich and quantify the labeled protein.

1.4.2 Click chemistry-based design concepts

The versatility and ease in synthesis of activity-based probes in both protein imaging and/or isolation can be achieved by the use of "clickable" handles to incorporate reporter groups postlabeling (**Fig. 1.17a**). In these cases, the probe and the tag are synthesized with a complementary pair of orthogonally reactive chemical entities. Staudinger ligation, which covalently links a phosphane and azido groups or the Cu(I) catalyzed 1,3-dipolar cycloaddition between an alkyne and an azide are examples of such click chemistry approaches. A desirable feature of this method is that it minimizes the size of the ABP—the incorporation of large fluorophores or tags at the outset can inhibit the delivery of the probe into the cell, such a two-step labeling approach could thus could provide the desirable utility for studying biologically relevant systems in their native state.

Cravatt and co-workers applied the principles of ABPP to develop a general method for accelerated target discovery in small-molecule cell-based screens[45]. They designed a natural products-inspired library with two features to assist target identification: (a) an electrophilic spiroepoxide to covalently modify cellular target and (b) an alkyne for the visualization of the target protein via click chemistry. The spiroepoxide library was screened for anti-proliferation activity against the invasive human breast cancer cell line MDA-MB-231. One compound from the probe library, MJE3, was identified to have the best inhibition on proliferation (**Fig. 1.17b**). On analysis of the in situ reactivity profile, a single protein was found to be uniquely labeled by MJE3 and this protein was identified as glycolytic enzyme phosphoglycerate mutase B.

In another recent report, Cravatt and co-workers devised ABPs targeting cytochrome P450 super family, which are of great pharmaceutical interests[46]. The probe was designed

to have a 2-ethylnapthalene unit (know inhibitor of cytochrome P450) tagged with an alkyne handle. The probe was catalytically oxidized to a reactive ketene, which then reacted with a nuclueophilic residue in the enzyme. Subsequently it was clicked with a rhodamine/biotin reporter units tagged with azide. The strategy was successful in profiling the drug-P450 interactions both *in vitro* and *in vivo* (**Fig. 1.17c**).

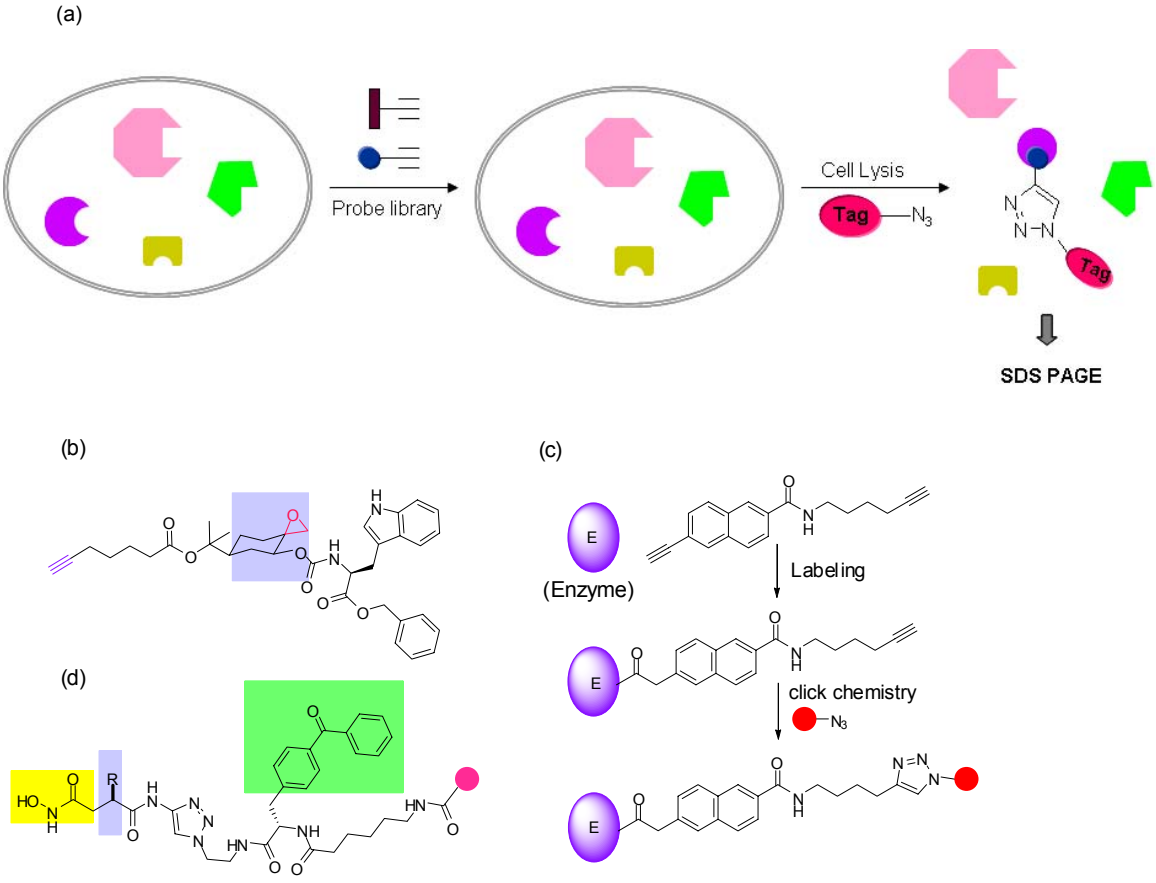


Fig 1.17: **(a)** 'Click' based two-step labeling approach **(b)** MJE3, natural product inspired activity-based probe **(c)** Labeling of cytochrome P450 **(d)** Labeling of metalloproteases.

Our group utilized click chemistry for the modular synthesis of ABPs targeting Matrix metallo-proteases (MMPs) which are laborious to synthesize individually[47]. This enabled the rapid assembly of a panel of hydroxamate-based trifunctional probes that each contained a benzophenone unit (photocross-linker) and a rhodamine tag (fluorophore).

The P_1' position of the ABPs was diversified with twelve natural and unnatural amino acids, which rendered characteristic fingerprints when profiled against a panel of several different metalloenzymes (**Fig. 1.17d**).

1.4.3 Imaging of Enzyme activity

Cravatt and co-workers have developed cell-permeable quenched activity-based probes (qABP) that can be used for direct imaging of papain-family cysteine protease activity in living human cells[48]. The probes consist of a peptide fragment attached to a reactive unit (acyloxymethyl ketone), fluorophore (rhodamine) and a quencher (BODIPY). The probe becomes fluorescent upon activity-dependent covalent modification of protease target. This probe is used to monitor real-time protease activity in living cells (**Fig. 1.18**).

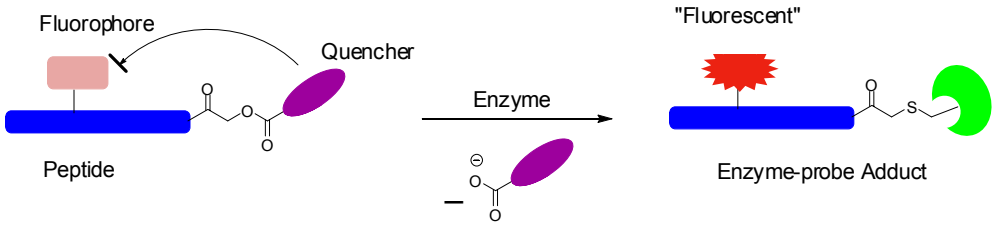


Fig. 1.18 Activity-dependent labeling of cysteine protease by qABP

Our lab has developed a handful of activity-based probes which is discussed in **chapter 5**. Other high-throughput activity-based profiling techniques/platform like microarray[49], 'expression display'[50] for the genome wide identification of enzyme activities were also developed in our lab.

1.5 Bioimaging

Activity-based probes has proven to be a vital tool in identifying the protein classes, the next challenging task lies in elucidating how such proteins performs their functional roles in their native cellular milieu. Visualizing proteins in living cells is essential for a better understanding of their structure and functions. With the advent of fluorescent proteins and other molecular labels, cell biologists and protein chemists have witnessed a better understanding of intracellular proteins and subcellular events. Together with advances in fluorescence and confocal laser scanning microscopies[51], 3-dimensional imagings of cellular events and protein dynamics have become possible in real time[52]. Traditionally, proteins were labeled *in vitro* with fluorescent and other molecular probes[53], then transferred into live cells and monitored, in real time, using advanced imaging techniques. Recent advances in genetic engineering have made it possible to directly generate fluorescent proteins in living cells or even in live animals by fusion of fluorescent proteins such as GFP (green fluorescent protein) to the protein of interest[54]. However, the main drawbacks of GFP-like proteins include their large sizes, obligate oligomerization, and sometimes slow or incomplete maturation. I have discussed few imaging techniques using small molecules in the below paragraphs. A good labeling strategy should ideally satisfy the following criteria: (1) possesses a high signal-to-noise ratio (i.e. high specificity for target protein); (2) upholds the integrity of the labeled protein; (3) does not interfere with the biochemical functions or cellular localization of the labeled protein and (4) has minimal perturbation to the normal cellular processes.

One of the earliest methods for site-specific labeling of recombinant proteins with small organic molecules within live cells was developed by Tsien *et al*.[55]. This method exploits

the well-known specificity of organoarsenicals with pairs of thiols. A short peptide sequence CCXXCC (in which X is a non-cysteine amino acid) was genetically fused to the protein of interest, which was subsequently recognized and labeled by a cell permeable fluorescein derivative called FlAsH (**Fig. 1.19a**). FlAsH contains two As(III) substituents conjugated to ethanedithiol (EDT). The free FlAsH label is virtually non-fluorescent, but acquiring a 50,000-fold increase in fluorescence upon binding to the CCXXCC motif.

Covalent Labeling Using hAGT Fusions strategy was developed by Keppler *et al.* whereby specific covalent labeling with a small molecule was achieved through the human DNA repair protein O6-alkylguanine-DNA alkyltransferase (hAGT)[56] (**Fig. 1.19c**). The native biological role of hAGT is to irreversibly transfer the alkyl group from O6-alkylguanine-DNA to one of its reactive cysteine residues, to aid in the repair of the alkylated DNA[57]. hAGT also reacts with the cell-permeable nucleobase O6-benzylguanine (BG) and BG derivatives substituted at the 4-position of the benzyl ring. By exploiting this chemistry, a hAGT fusion protein expressed in live cells was labeled with BG derivatives, BGBT and BGAF, which contain biotin and fluorescein, respectively.

Vogel *et al* reported a generic method for the site-selective and reversible labeling of membrane proteins containing a polyhistidine sequence in live cells with small molecular probes[58] (**Fig. 1.19b**). The labeling strategy was based on the well-known interaction between polyhistidine sequences and the Ni^{2+}-NTA moiety[59], in which a small organic fluorophores conjugated to nitrilotriacetate (NTA) moiety were used. The structure, as well as the plasma membrane distribution, of the ionotropic 5-hydroxytryptamine

serotonin receptor ($5HT_3R$) was studied using this approach. Our own contribution to the field of bioimaging involves a novel strategy for site-specific covalent labeling of proteins *in vivo*[60] by taking advantage of the chemoselective reaction of native chemical ligation[61], which is explained in detail in the last chapter of the thesis.

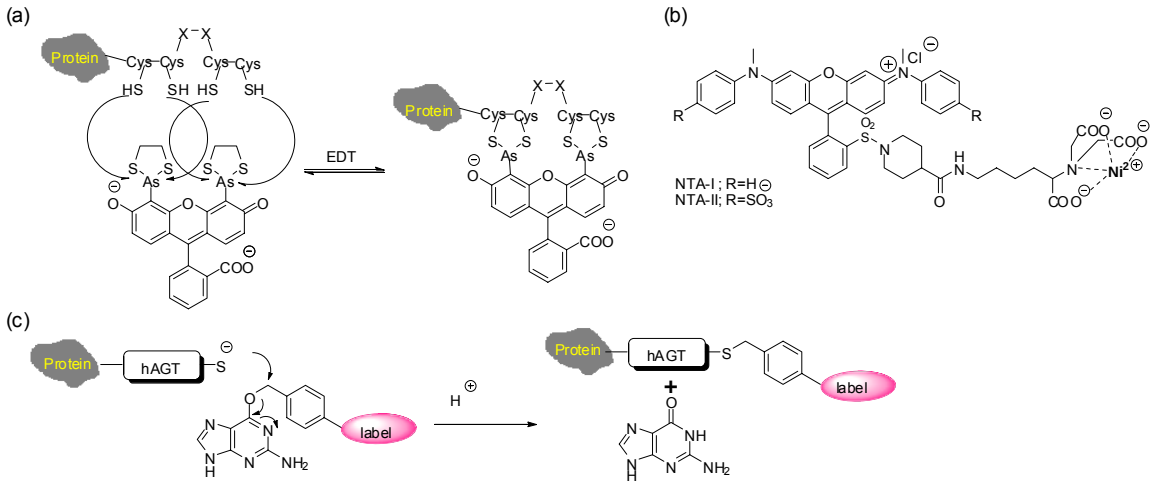


Fig. 1.19 (a) Site-specific labeling of tetracysteine motif by FlAsH **(b)** NTA probe coordinating to Ni (II) ion. **(c)** Labeling of hAGT fusion protein with BG derivative

.

Chapter 2

High-Throughput Assembly of Protein Tyrosine Phosphatases (PTPs) Inhibitors Using "Click Chemistry"

2.1 Summary

Click chemistry coupled with *in situ* screening is the simplest and versatile of all the fragment-based drug discovery approaches. This chapter summarises the development of a 66-member bidentate inhibitors of PTPs, by employing click chemistry for the first time, in the fragment based assembly of PTP1B. Screening of the library against a panel of different PTPs and non-PTPs revealed a specific inhibitor of PTP1B (IC_{50} = 4.7 μM). The second part of this chapter narrates the successful traceless solid-phase synthesis of 325-member azide library suitable for direct click chemistry applications. The utility of the library is demonstrated with its subsequent high-throughput "click" assembly of 3250-member library of bidentate inhibitors targeting various PTPs.

2.2 Introduction

2.2.1 Fragment-based drug discovery of PTP inhibitors

Fragment-based drug discovery approach is currently predominant in pharmaceutical industries[62]. The approach aims at identifying weekly active ligands (K_D ~ 100-1000 μM)

and translates them into low nanomolar inhibitors against targets of low or borderline druggability targets using a variety of design strategies. Currently the most common strategy to assemble the compounds includes: (1) SAR by NMR/X-ray/MS strategy, (2) Tethering approach by Sunesis, (3) in situ screening method and recently Ellman's SAS method. In SAR by NMR method, small organic molecules that bind to proximal subsites of a protein are identified, optimized and linked together to produce compounds with high binding affinity. Scientists at Abbott laboratories have identified salicylic acid based site 2 binding ligand (K_d = 1.2 mM) and isoxazole carboxylic acid (K_d = 800 μM) as catalytic site binder of PTP1B through an NMR based screening of over 10,000 compounds (**Fig. 2.1a**).

Scientists at Sunesis came up with a platform called "Tethering" which allows for the identification of small-molecule fragments that bind to specific regions of a protein target. These fragments can then be extended, combined with other molecule or combined with one another to provide high-affinity drug leads. Standard covalent tethering uses a cysteine residue to capture disulfide-containing fragments that binds to a site near to the active site on the protein and mass spectroscopy is generally used to identify the fragments (**Fig. 2.1b**). They extended this technology to "breakaway tethering", which overcomes the difficulty of introducing a cysteine residue in the highly conserved active site of PTP1B and used to discovery novel ligands that binds to the active-site of PTP1B[63].

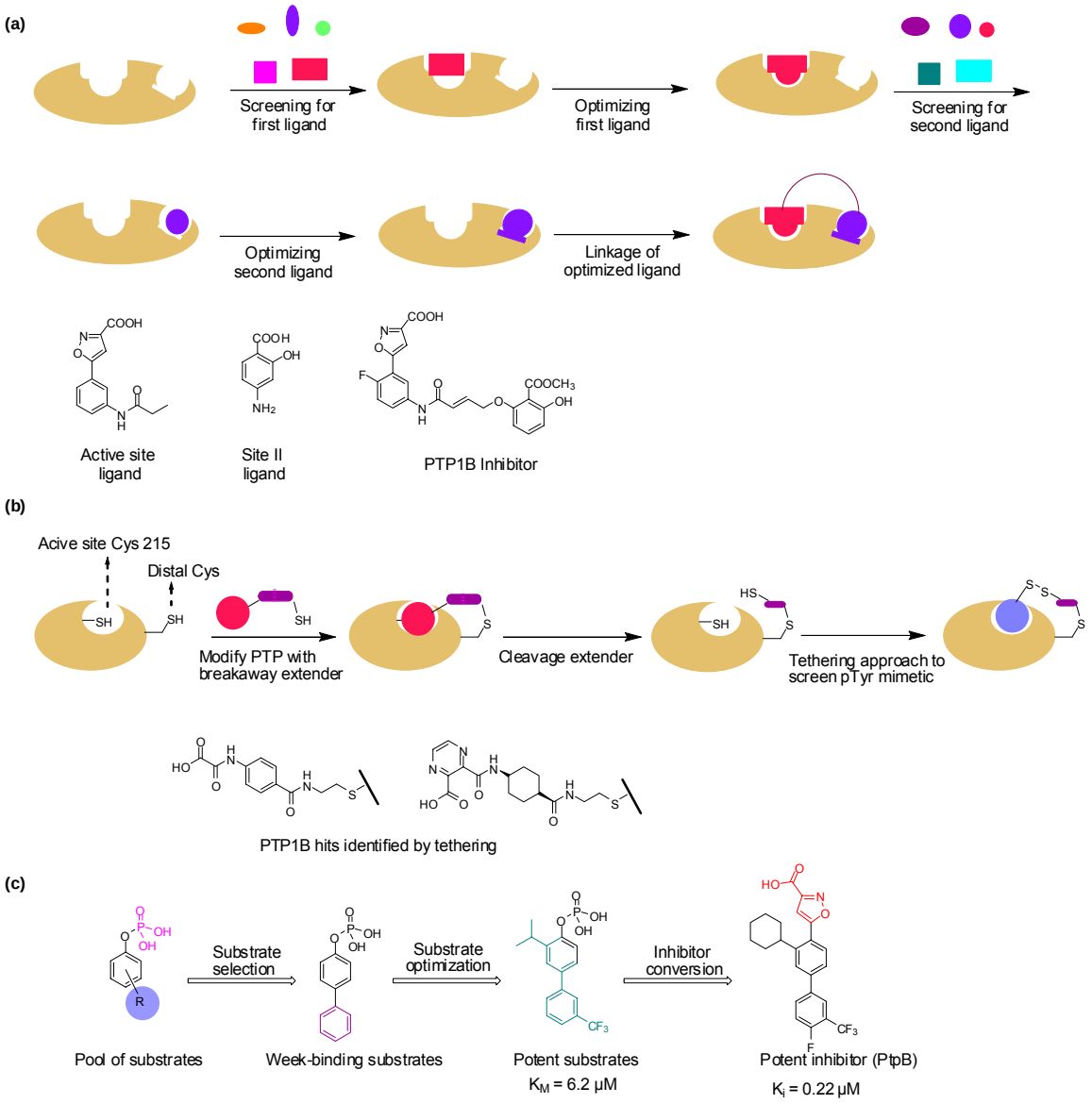


Fig. 2.1 Different fragment-based approaches to discovery PTP inhibitors **(a)** SAR by NMR method **(b)** 'Tethering' approach **(c)** Substrate activity screening method

Recently, Ellman's group devised a new fragment-based approach called substrate activity screening (SAS) to identify selective and low-molecular-weight PtpB inhibitors with submicromolar inhibitory activity (**Fig. 2.1c**). Their approach consists of 3 steps: (1) 140-member library of diverse *O*-aryl phosphates were screened to identify phosphate

substrates using a simple spectrophotometric-based assay; (2) optimization of the O-aryl phosphate substrates by analogue synthesis and evaluation; and (3) replacement of the phosphate group with known isosteres. By this new approach they were able to fish-out the most potent PtpB inhibitor reported to date (K_i = 0.22 μM) with low molecular weight and good selectivity against a panel of other PTPs[64].

2.2.2 Click Chemistry as a High-throughput tool

Click chemistry has emerged as an integral part of the drug discovery pipeline by providing a high-throughput amenable chemical reaction platform for compound synthesis[65, 66]. Unlike other fragment-based approaches, click chemistry-based strategies require neither sophisticated instruments nor mutations in the target proteins, and at the same time enable the exploration of N^2 possibilities with N + N chemical entities. This is particularly useful for small-molecule lead discovery against enzymes (and other proteins) that possess an extended active site (as in the case of most proteases) or multiple binding pockets[67] (as in the case of most kinases and phosphatases) (**Fig. 2.2**).

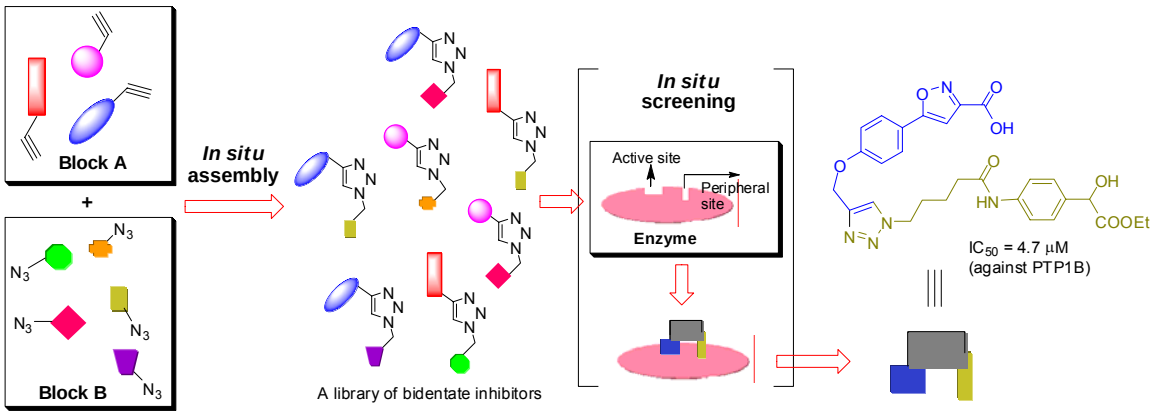


Fig 2.2 Click chemistry and *in situ* screening of inhibitor discovery

2.2.3 Bidentate inhibitors against Protein Tyrosine Phosphatases 1B (PTP1B)

PTPs are the key signaling enzymes with plays a vital role in myriad of cellular processes[4]. There are ~40 PTPs in the human body and among the various PTPs, PTP1B is the prototype and recently identified to be a key player in diabetes and obesity and PTP1B is recognized as a druggable proteome[68]. Active research has been going on to discover potent and specific small molecule inhibitors against PTP1B. However the main challenge lies in developing a specific inhibitor against PTP1B, because all the PTP's (including PTP1B) active sites are highly conserved and have high sequence homology. PTPs like LAR, CD45, SHP-2, cdc45 and T-cell PTP (TCPTP) share 51, 48, 52, 44 and 80% homology in the catalytic domains to PTP1B which results in a challenging task to inhibit PTP1B selectively.

Recently, Zhang *et al* discovered the presence of a secondary aryl-phosphate binding site near the active site of PTP1B, thus shedding lights on the development of bidentate inhibitors which might impart both potency and specificity against the enzyme[69] (**Fig. 2.3**). Based on this model, researchers at Abbott laboratories recently discovered cell-permeable, micromolar bidentate inhibitors containing a core N-phenyloxamic acid mimic which, upon testing in PTP1B-expressing COS-7 cells, showed much greater cellular activities (in terms of inhibition and selectivity) and pharmacokinetic properties than other inhibitors possessing potent inhibitory activity *in vitro* but suffering from low cell permeability and selectivity *in vivo*.[70,71] The key to their success is the use of NMR-based fragment assembly approach which greatly facilitated the rational improvement of lead compounds[72].

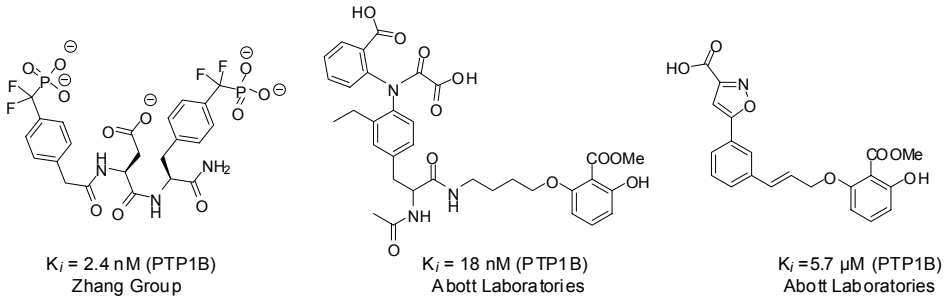


Fig. 2.3 Bidentate inhibitors against PTP1B

2.2.4 PTP1B bidentate inhibitor design

We designed a total of 5 alkyne functionalized isoxazole-based warhead which can binds to the PTP1B active sites and we designed and synthesized 14 different azide functionalized aromatic blocks which can binds with the peripheral sites. We explored the in situ "Click Chemistry" to assemble these fragments. In short, our inhibitor design encompasses the following (**Fig 2.4a**).

(1) Each member is made of a modular, bidentate structure containing both a core and a peripheral group for potential binding to the enzyme (**Fig. 2.4b**);

(2) Whenever possible, each of the two (core and peripheral) groups should possess optimized pharmacological properties.

(3) An alkyl linker (with optimized chain length) which links the active site binder with the peripheral binding group

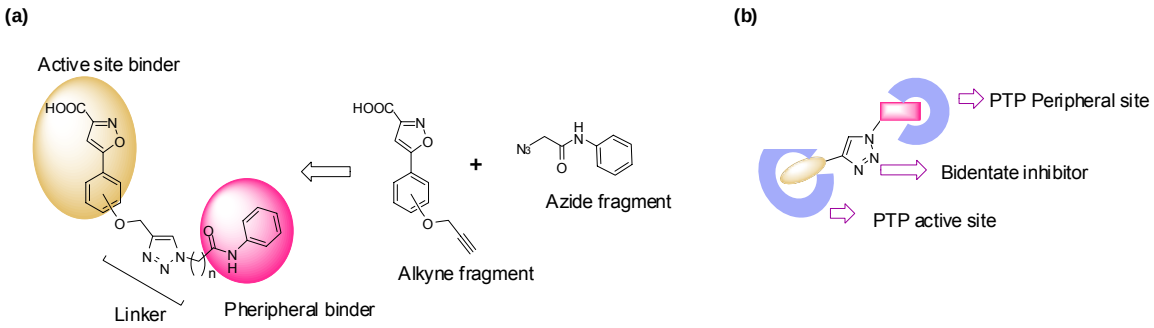


Fig. 2.4 (a) Click assembly of PTP1B inhibitors **(b)** Mode of binding of the PTP1B bidentate inhibitor

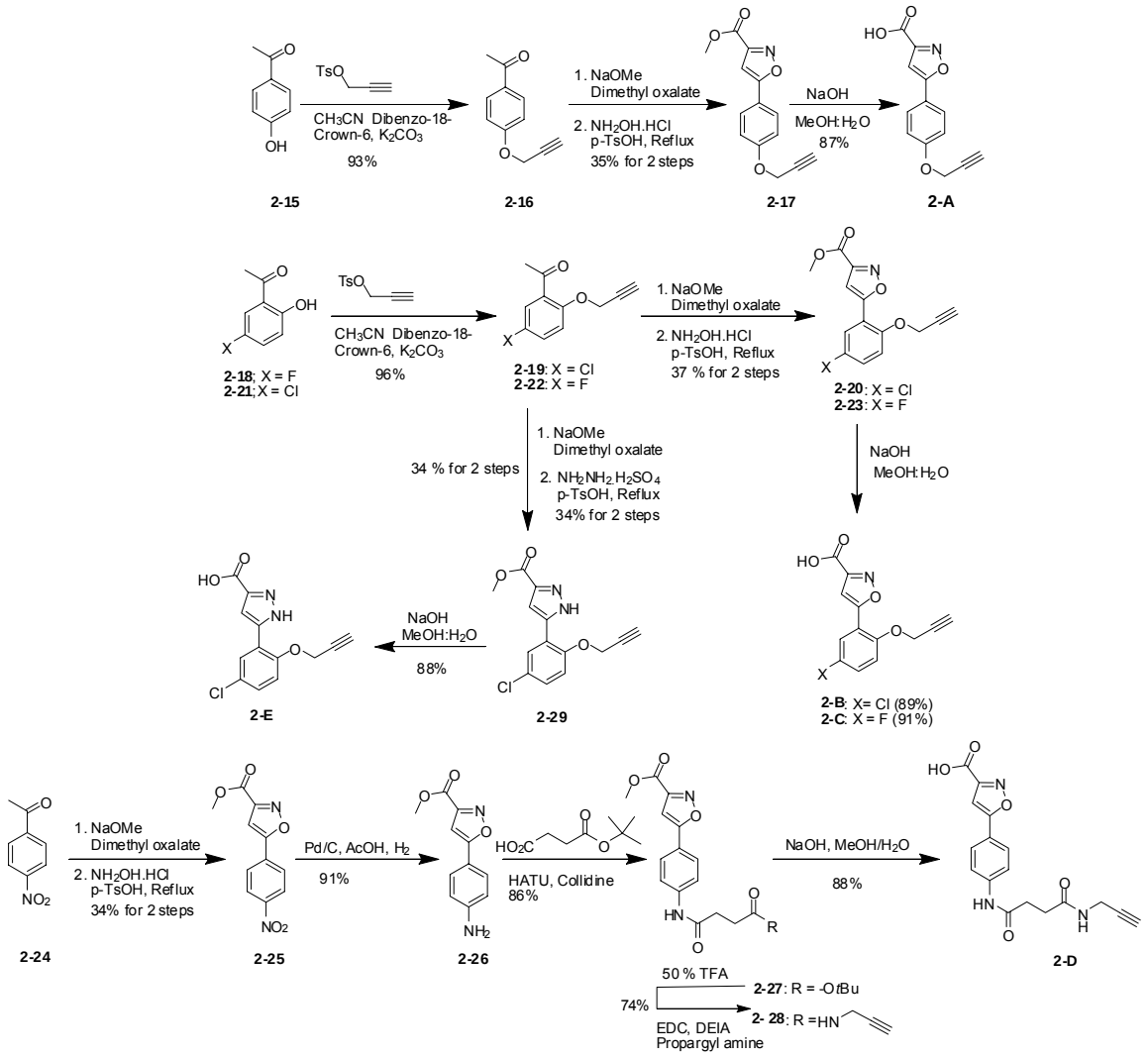
2.3 Results and Discussion

2.3.1 Chemical Synthesis of the inhibitor library

Synthesis of the Alkyne fragments:

Compound **2-A** was synthesized from the commercially available 4-hydroxyacetophenone, **2-15**, which upon refluxing with propargyl tosylate in the presence of K_2CO_3 and dibenzo-18-crown-6, afforded **2-16** with excellent yield (93%). Followed by condensation between **2-16** and dimethyloxalate in the presence of NaOMe, followed by cyclization of the resulting product in the presence of hyroxylamine, gave the isoxazole carboxylic methylester, **2-17**, in modest yield (35 % for 2 steps). Similarly starting from 5'-chloro- or 5'-fluoro-2'-hydroxyacetophenone, compounds **2-B** and **2-C** were synthesized(i.e. **2-18** or **2-21**). For the synthesis of compound **2-D**, commercially available 4-nitroacetophenone **2-24** was condensed with dimethyloxalate followed by cyclization to afford **2-25** (36% overall yield). Subsequent reduction of the nitro group in **2-25** with 10% Pd-C in the presence of H_2 gave **2-26** (91%), which upon coupling with mono-Boc protected succinic acid, **2-36**, generated **2-27**. Next, deprotection of the Boc group by 50 % TFA in dichloromethane followed by coupling with propargylamine

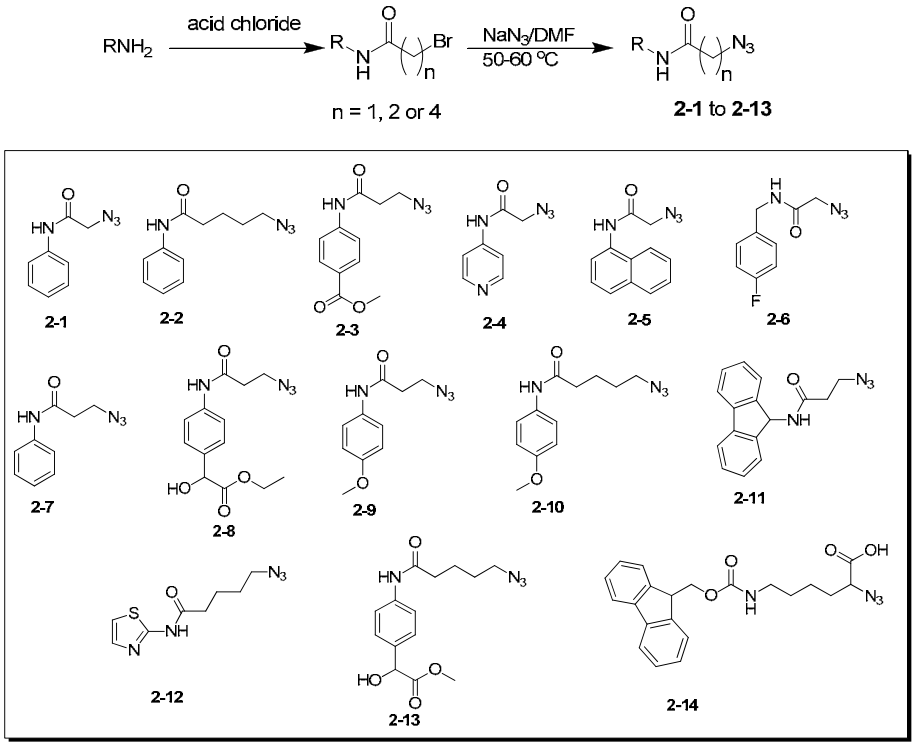
afforded **2-28**. The methyl ester in **2-28** was subsequently hydrolyzed with NaOH to afford compound **2-D** in 88% yield. In the case of **2-E**, condensation of **2-19** with dimethyloxalate followed by direct treatment of the resulting product with hydrazine hydrate in the presence of *p*-TsOH generated the pyrazole carboxylic methylester, **2-29**, which upon hydrolytic cleavage with NaOH afforded compound **2-E** in 88% yield, (**Scheme 2.1**).



Scheme 2.1 Synthesis of the alkyne building blocks

Synthesis of the azide fragments:

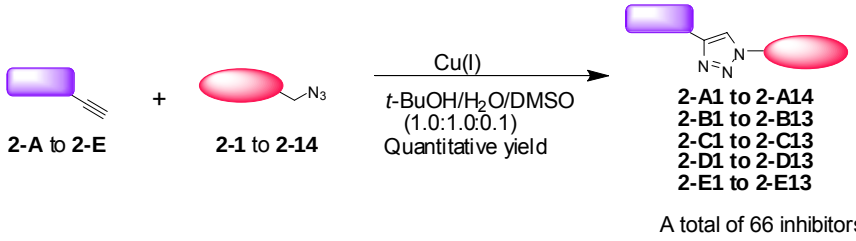
The 14 different aromatic azides were synthesized using a two/three-step procedure (**Scheme 2.2**). 2-bromoacetyl chloride and 3-bromopropionyl chloride were purchased commercially, whilst 5-bromovaleryl chloride was synthesized from 5-bromovaleric acid, using thionyl chloride. Subsequently, 9 different aromatic amines, representing different polarity (benzyl, phenyl, napthyl, heterocycles) and subsituents (halogens, esters, ethers, alcohols), were acylated with one of the above three bromoalkylacyl chlorides, followed by an S_N2 substitution reaction with sodium azide in DMF to generate the corresponding azides. Compound **2-14** was synthesized from the corresponding amino acid (i.e. lysine).



Scheme 2.2 Synthesis of Azide building blocks

Assembly of the Library

The Cu (I)-catalyzed ligation of alkyne blocks and azide blocks was next carried out. It was found that a mixed solvent system containing *t*-BuOH, water and DMSO (1:1:0.1) enabled both the alkynes and the azides to be completely dissolved, and at the same time the 1,3-dipolar coupling be carried out with extremely high efficiency. More importantly, the products could be taken directly without further purifications. 5 different alkynes were clicked with 13 different azides [(5 x 12) + (1 x 1)] to afford a 66 member library (**Scheme 2.3**). LC-MS analysis of all 66 coupling reactions indicated the complete consumption of the alkynes and quantitative formation of the triazole products in most cases.

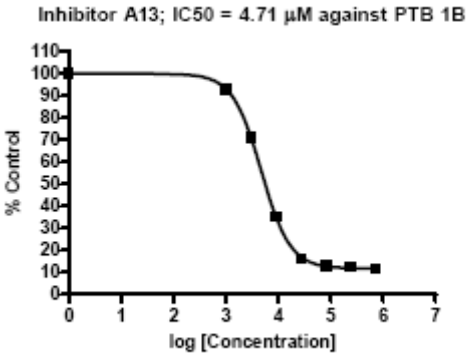


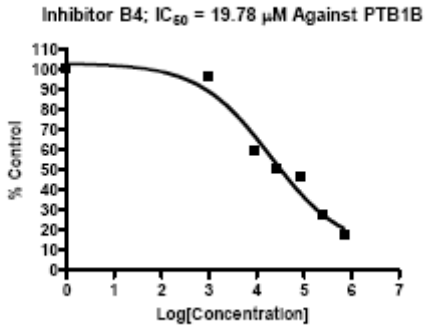
Scheme 2.3 Click assembly of PTP inhibitors

2.3.2 Results and conclusion of the screening experiments

The 66-member library was next screened directly against 6 different phosphatases, including 4 PTPs (i.e., PTB1B, TCPTP, YOP and LAR), a dual-specific phosphatase (i.e. Lambda PPase) and serine/threonine phosphatases (i.e. PP1). Fluorescence-based phosphatase assay was used throughout[73]. A majority of the library members showed some degree of inhibition against three of the four PTPs, namely PTB1B, TCPTP and YOP. This is expected when one considers that the library was purposefully designed to contain a core group which structurally resembles the original Abbott's inhibitor. On the

other hand, none of the library members showed any inhibition against LAR, λPPase and PP1 until at very high concentrations (> 450 μM of inhibitor). For LAR, a protein tyrosine phosphatase, this is somewhat surprising. On the other hand, this may also underscore the feasibility of our strategy for potential discovery of inhibitors against specific PTPs with high potency and specificity. Compounds **2-A13**, **2-C10**, **2-B1**, **2-B11**, **2-B4**, **2-A5** were identified to be the relatively potent ones against PTB1B, TCPTP, YOP, LAR, λPPase, and PP1, respectively. These six compounds were further investigated by the measurement of their IC_{50} values. The IC_{50} values of the six compounds against all six enzymes were shown in previous section. The IC_{50} values were obtained using dose-dependent reactions by varying the concentrations of the inhibitor, under the same enzyme concentration. Representative results are shown in **Fig. 2.5** and the rest are given in the experimental section.





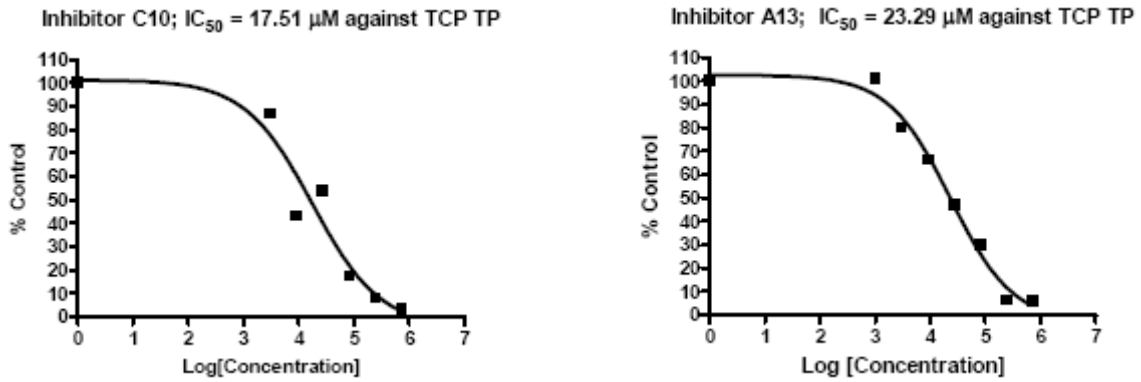


Fig. 2.5 IC_{50} graphs for selected inhibitors against PTP1B and TCP TP

Among the six hits (**2-A13**, **2-C10**, **2-B1**, **2-B11**, **2-B4** & **2-A5**) identified, one of which, e.g. **2-A13**, a potential potent and specific PTP1B inhibitor, and further characterized. The structures of the 6 inhibitors are shown in **Fig. 2.6**. As shown in **Table 2.1**, compound **2-A13** indeed inhibited PTP1B with an IC_{50} of 4.7 µM, which is comparable to that of the original Abbott's inhibitor. More importantly, it is approximately 5 and 25 times more selective towards PTP1B than the other two PTPs, TCPTP and YOP, respectively, and is > 100 times more potent than against LAR, λPPase and PP1. Another noticeable hit is **2-C10**, which showed good inhibition towards TCPTP (IC_{50} = 17.5 µM) with some degree of selectivity over other phosphatases.

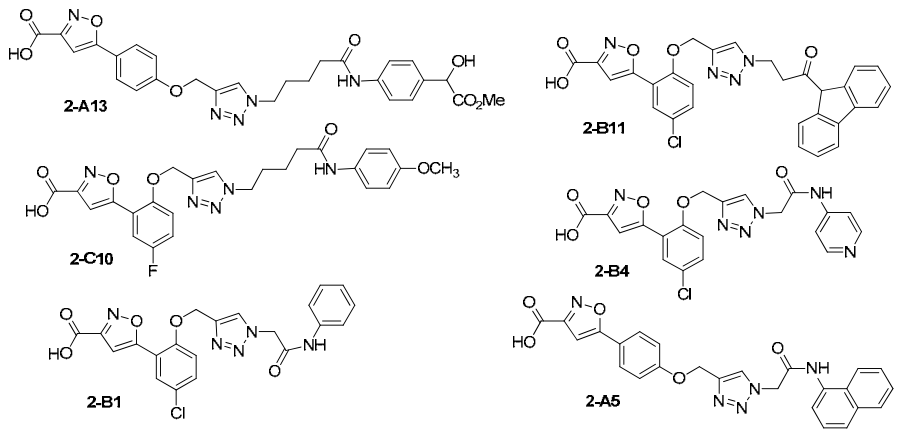


Fig. 2.6 Structures of the potent inhibitors identified from the screening.

Inhibitor	PTP1B	TCPTP	YOP	LAR	λPPase	PP1
2-A13	4.7	23.3	120.1	>450	>500	>750
2-C10	58.1	17.5	24.1	>450	>500	>750
2-B1	17.2	25.5	23.9	>450	>500	>750
2-B11	29.6	21.7	82.2	>450	>500	>750
2-B4	19.8	49.4	176.8	>450	>500	>750
2-A5	34.5	64.5	74.6	>450	>500	>750

Table 2.1 IC_{50} (in μM) of 6 selected inhibitors.

2.4 High-throughput synthesis of 3250-member PTP inhibitor library

2.4.1 Traceless solid-phase synthesis of 325 azide fragments

Click chemistry-based screening method is highly advantageous in identifying enzyme inhibitors but it remains to be truly high-throughput amenable due to the lack of a robust and efficient chemical strategy that give access to large, diverse libraries of azide functionalized enzyme binders using readily available building blocks (like alcohols or amines). Standard methods of preparing azides[74] involves the solution phase synthesis and purification of the individual compounds by column purification, therefore it is time-consuming and low-throughput. We report a truly high-throughput amenable, solid-phase strategy for the synthesis of high-quality azides which, without purification, could be

used for “click assembly” and direct in situ screening of enzyme inhibitors. Our strategy made use of the commercially available PL-FMP resin (aldehyde resin) on which an immobilized 4-formyl-3-methoxyphenoxy moiety efficiently captured a variety of primary amines via reductive amination and converted them to the corresponding secondary amines which were subsequently acylated. Following TFA cleavage, the desired products were obtained in uniformly high quality (**Fig. 2.7**). There are numerous advantages in this method: (1) it is a traceless solid-phase strategy, allowing the use of as-is starting materials (i.e. same as solution-phase methods) as building blocks and is easily amenable to high-throughput synthesis (i.e. IRORI™ split-and-pool sorting technology) of a large number of azides. (2) It is highly robust, capable of making a variety of azides from different functionalized building blocks (i.e. amines, acids, sulfonyl chloride, and chloroformate) with suitable linkers; (3) it gives rise to products with extremely high purity, thus allowing them to be used directly for “click” assembly followed by in situ enzyme screening.

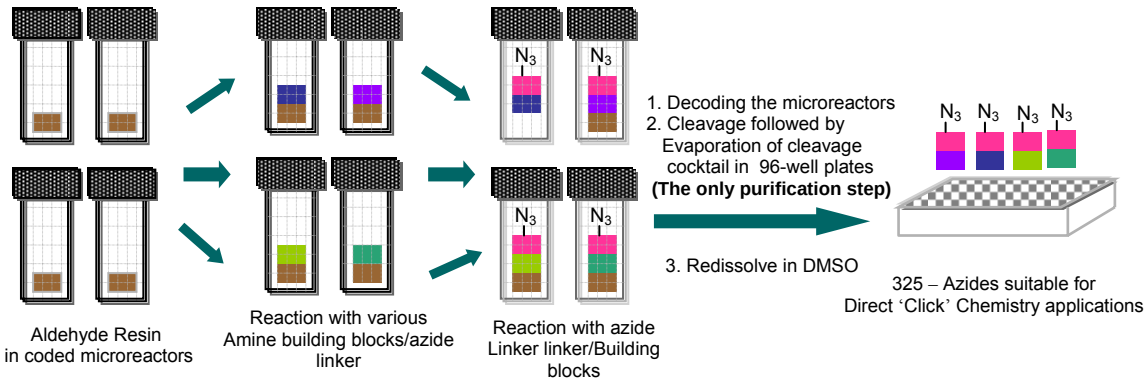
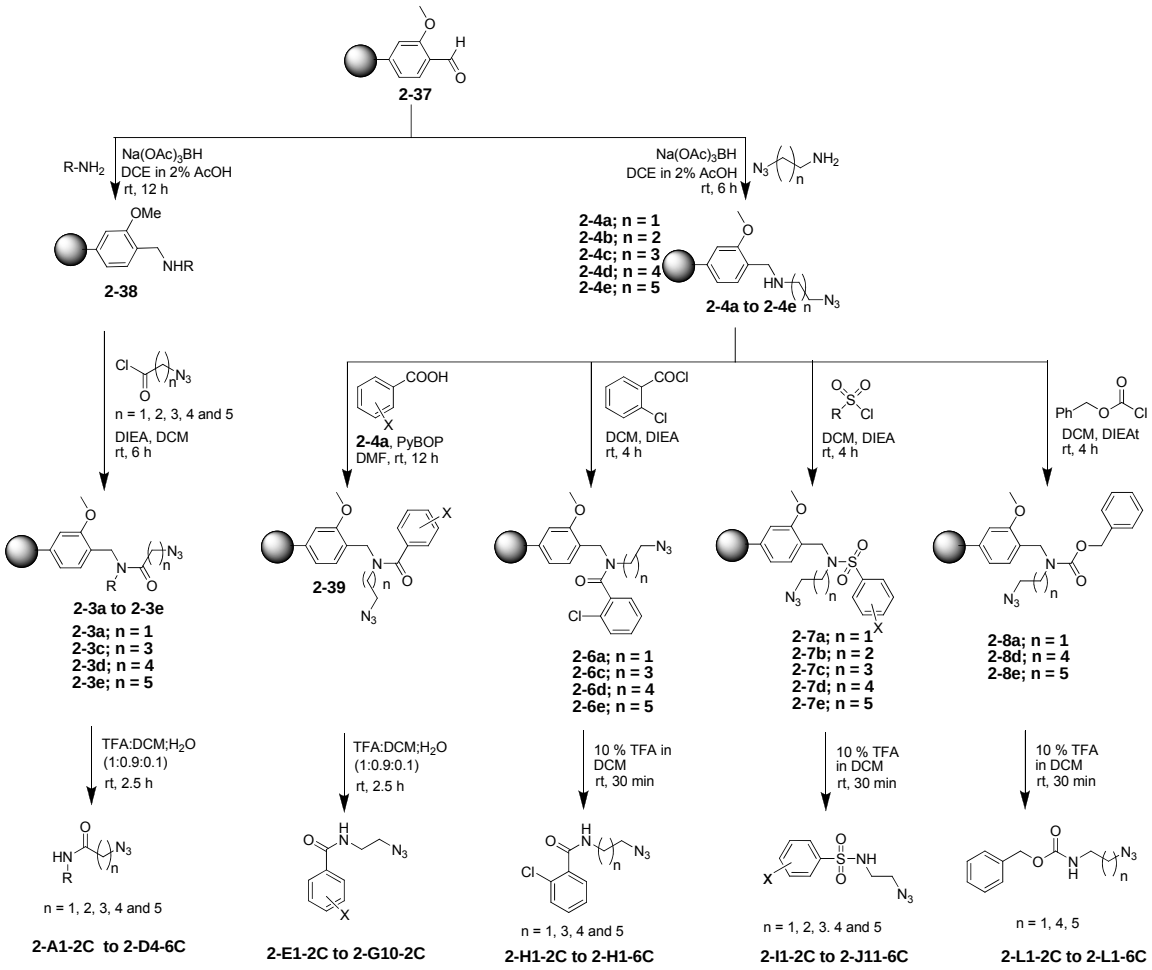


Fig. 2.7 Schematic representation of high-throughput azide synthesis

2.4.2 Chemical synthesis of the azide fragments

To demonstrate our strategy, a 325-member azide library was synthesized as shown in **Scheme 2.4**. Depending on the choice of building blocks and linkers, route A or B was taken. In route A, 40 different commercially available aromatic amine building blocks were treated with the PL-FMP aldehyde resin, **2-37**, in the presence of $Na(OAc)_3BH$/2 % glacial acetic acid, to give the corresponding secondary amines, **2-38**. The reductive amination reaction proceeded smoothly with a variety of aromatic amines bearing different substitutents such as -F, -Cl, -OR, $-SCH_3$, $-CO_2R$ and -R. Benzyl, napthyl and anthracenyl amines too underwent reductive amination smoothly. Subsequently, treatment with azidoalkynoyl chlorides, furnished the N-acylated products, **2-3a** to **2-3e**. Other coupling conditions attempted (e.g. DIC/EDC/HATU/PyBOP) failed to give the desired products in sufficient purity. In route B, we first carried out reductive amination with azide-functionalized amine linkers to give **2-4a** to **2-4e**. Subsequently, the resulting resins were treated with different commercial available building blocks, including acid chloriders, aryl acids, sulfonyl chlorides and chloroformates, giving **2-6a** to **2-6e**; **2-39**, **2-7a** to **2-7e**, and **2-8a** to **2-8e** respectively. We found that coupling using acid chlorides, sulfonyl chlorides and chloroformates was excellent as expected. With aryl acid building blocks, however, the PyBOP/HOAt coupling was the method of choice, providing complete coupling to give **2-39**. Other coupling conditions were tried but failed to give satisfactory results. We also found building blocks bearing different functionalities such as –F, -Cl, -Br, -R, -OH, -OCH3, $-CF_3$, -CN, -OH, $-NO_2$ and -vinyl were tolerated. Finally, cleavage of the resins with optimized TFA cocktails gave the desired products, which, upon concentration in vacuo, analysis and characterizations by LC-MS/NMR,

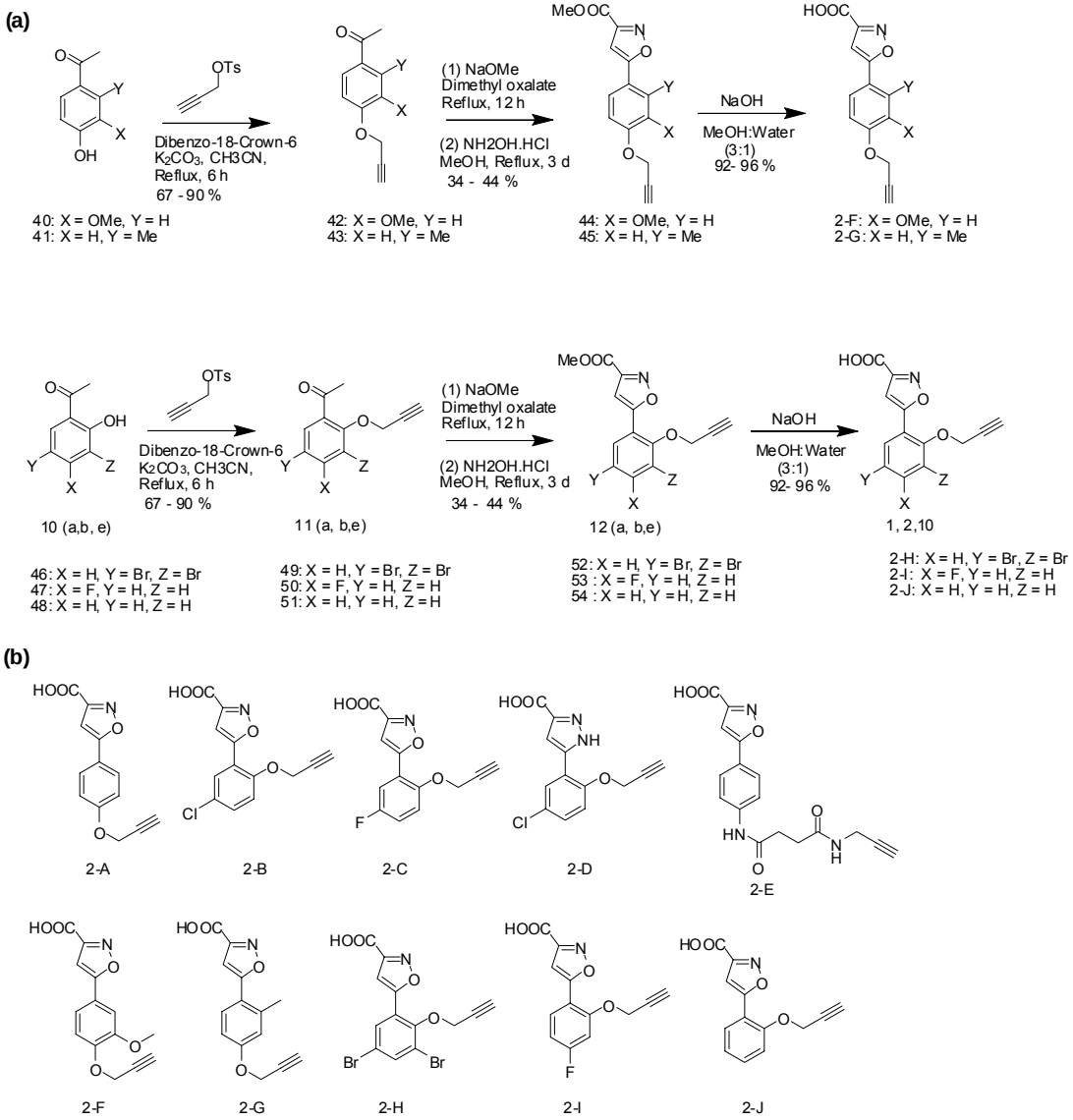
showed uniformly high purity (> 95% in most cases). In total, 198 amide-bearing azides (i.e. **2-A1-2C** to **2-D4-6C**, **2-E1-2C** to **2-G10-2C**, **2-H1-2C** to **2-H1-6C**), 115 different sulfonamide-containing azides (i.e. **2-I1-2C** to **2-J11-6C**), and 3 carbamate-derivatized azides (i.e. **2-L1-2C** to **2-L1-6C**) were successfully prepared. Stock solutions of these azides were prepared in uniform concentration with DMSO in 96-deep well plates and used directly, without any purification, for subsequent *in situ* click assembly followed by enzymatic screening.



Scheme 2.4 Traceless solid-phase synthesis of azide libraries

2.4.3 Chemical synthesis of the 3250-member click library

Five different alkynes (set II) functionalized with PTP warheads were synthesized as shown in **Scheme 2.5**



Scheme 2.5 (a) Synthesis of alkynes set-II (b) Structure of the 10different alkynes used

Having successfully synthesized a variety of 325 azides and 10 different alkynes in total, next we assembled them in a high-throughput fashion by employing the Cu(I) catalyzed alkyne-azide 'click' ligation (**Fig. 2.8**).

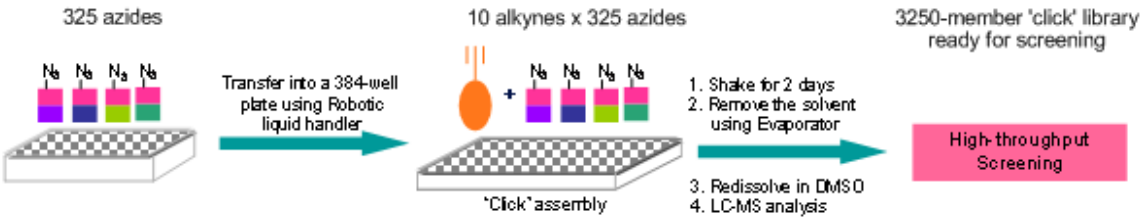


Fig. 2.8 Schematic representation of 3250-member 'click' assembly

At first we optimized the ligation by attempting different reaction conditions and solvents (**Table 2.2**). In our hands $CuSO_4$/Sodium ascorbate system in either t-BuOH or 1,2-dichloroethane(DCE) or dichloromethane (DCM) as co-solvent in water worked well. However, since chlorinated solvents were not compatible with 384-well plates due to their easy-volatile nature, we started the library synthesis using water/t-BuOH as solvent.

No.	Catalyst	Additive	Solvent	Yield	Purity
1*	$CuSO_4.5H_2O$	Sodium ascorbate	t-BuOH	~100	>95
2	$CuSO_4.5H_2O$	Sodium ascorbate	DMSO	>95	>95
3	$CuSO_4.5H_2O$	Sodium ascorbate	DCM[§]	~100	~100
4	$CuSO_4.5H_2O$	Sodium ascorbate	1,2-DCE[§]	~100	~100
5	CuI	Acetonitrile, Pyridine	t-BuOH	>90	80
6	CuI	Acetonitrile, pyridine	DMSO	>95	90

*Best reaction condition in 384-well plate
[§]Easily evaporated from the 384-well reaction plate, need regular top-up of the solvent.

Table 2.2 Optimization conditions of the Click reaction

The alkyne solution in DMSO were dispensed to each well of the 384-well plate using a bulk-dispenser and the azides were dispensed directly from the 96-well plate to the 384-well plate with the aid of a robotic liquid handler, followed by the transfer of the catalyst mix ($CuSO_4$ and Sodium ascorbate solution in water) using the bulk dispenser. The plates were sealed using a silicon cap mat and shaken for 2 days after which the solvents were evaporated using a vacuum centrifuge and redissolved in DMSO and made into 50 mM stock plates and were analyzed and characterized by LC-MS. The limiting alkyne war-head was completely used-up and the desired triazole products were formed in quantitative yield and were taken directly for *in situ* screening.

2.5 Conclusion

In conclusion, we have successfully developed bidentate inhibitors of PTPs, by employing click chemistry for the first time, in the fragment based assembly of PTP1B inhibitors. A 66-member "click" library using five different alkyne-functionalized isoxazole-based active-site binders and 13 different azide-functionalized aromatic peripheral binders were synthesized in mixed water and t-BuOH solvent system using $CuSO_4$/Sodium ascorbate as the catalyst. Subsequent screening of the library again four PTPs (PTP1B, TCPTP, YOP and LAR), a dual specific phosphatase (PPase) and a serine/thronine phosphatase (PP1) revealed six potent inhibitors of the PTPs with one compound identified as a specific inhibitor of PTP1B (IC_{50} = 4.7 μM). None of the library members showed considerable inhibition against non-PTPs.

The second part of this chapter narrates the successful solid-phase synthesis of 325-member azide library for the first time via IRORI split-pool synthesis and subsequent

"click assembly." The method is highly robust and is highlighted by its simplicity and product purity. The utility of the library is demonstrated with its subsequent "click" synthesis with 10 different alkyne-functionalized isoxazole-based PTP active-site binders which gave access to 3250-member library of bidentate inhibitors targeting various PTPs. The entire operation was performed on 96/384 well plates and its truly high-throughput. With this new method, the whole process of inhibitor discovery, starting from the synthesis of building blocks to the identification of enzyme inhibitors, could now be completed in a very short time. We therefore envisage that this approach will provide a useful tool in the emerging field of "Catalomics."

Chapter 3

Solid-Phase Assembly and *in situ* screening of Protein Tyrosine Phosphatases inhibitors

3.1 Summary

This chapter summarises the development of a traceless solid-phase strategy to assemble PTP inhibitors using amide-bond formation reaction. The strategy is robust, giving high-quality products which are spectroscopically pure enough to be used for biological screening without any further purification. A 70-member combinatorial library of isoxazole-based bidentate inhibitors of PTPs were synthesized, which upon direct *in situ* screening revealed a potent inhibitor (K_i = 7.0 μM) against PTP1B.

3.2 Introduction

3.2.1 High-throughput amenable chemical reactions in inhibitor discovery

High-throughput study of enzymes is very crucial in the post-genomic era[75, 76]. One of the main challenges in the filed of high-throughput Enzymology is the development of high-throughput (HT) amenable chemical reactions that allow rapid synthesis of diverse chemical libraries for the interrogation of different classes of enzymes. One such reaction is the Cu (I)-catalyzed, 1,3-dipolar cycloaddition between an azide and an alkyne

fragment, Another class of reaction possessing similar qualities is the amide bond-forming reaction between an amine and a carboxylic acid using suitable activating/coupling reagents (like EDC, DCC, PyBOP, HATU, HBTU etc)[77]. The salient features of this reaction include its high yield, inoffensive by-products, compatible with microtiter plates, product stability etc. Numerous research groups have recently used this reaction for solution-phase, rapid assembly of small molecule inhibitors against a variety of enzymes including HIV proteases, β-aryl sulfotransferase, α-fucosidases and SARS-3CL protease[78]. The reaction is highly efficient, often generating the desired products in nearly quantitative yields, thus allowing *in situ* biological screening to be carried out directly in some cases, even in the presence of excessive starting materials, reagents or by-products (**Fig. 29a**). Apart from 1,2,3-triazole formation and amide bond-formation reactions, other high-throughput amenable chemical reactions like (1) Pictet-Spengler reaction; (2) N- and O-alkylation reactions; and (3) epoxide opening reaction have been adapted by various research groups in the discovery of enzyme inhibitors.

3.2.2 Limitations of Wong's *in situ* screening approach

Wong *et al* pioneered the microtiter-plate based reaction coupled with *in situ* screening in the discovery of enzyme inhibitors (**Fig. 3.1a**). Usually the reactions are done in microtiter-plate in microgram scale and directly taken for in situ screening without any further purifications or isolation of the product. The effectiveness of this method, however, had recently been challenged, as it was discovered that unexpected by-products may give rise to false results. In one typical example, Wong's group has attempted the synthesis of Lopinavir®-like compounds via amide bond-formation reaction between

various acids and amines in microtiter plates using HBTU as a coupling reagent. However, no desired products were formed and instead the intermediate benzotriazole ester was found to be responsible for the inhibitory activity towards SARS 3CL proteases[79] (**Fig. 3.1b**). To avoid such potential complications, we aim to develop solid-phase, amide bond-forming reactions using the same sets of starting materials as in the solution-phase

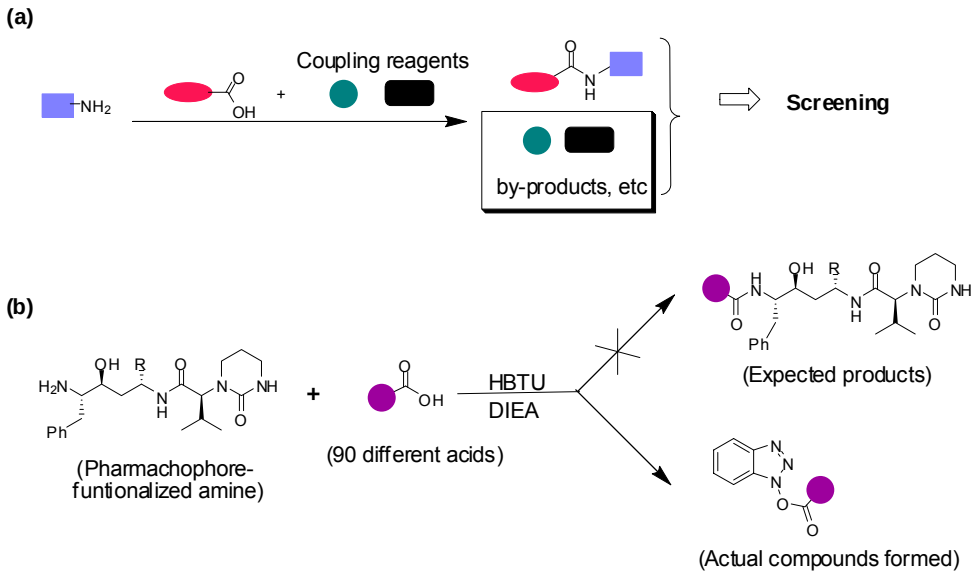


Fig. 3.1 (a) Wong's solution-phase and *in situ* screening approach **(b)** Stable benzotriazole by-products in amide-bond formation reaction.

3.2.3 Introduction to solid-phase combinatorial library

Combinatorial chemistry has emerged as a standard platform to generate large libraries of compounds for the drug discovery applications. Parallel and split solid-phase syntheses are highly useful in generating discrete and diverse sets of molecules respectively. In one specific example, Merck developed a new solid-phase strategy for the synthesis a library of 902 compounds of Indinavir® analogues (**Fig 3.2**)[80]. On

evaluation of these compounds, 2,6-dimethyl-4-hydroxy phenol was found to be a good replacement for the metabolically unstable aminoindanol moiety in Indinavir®.

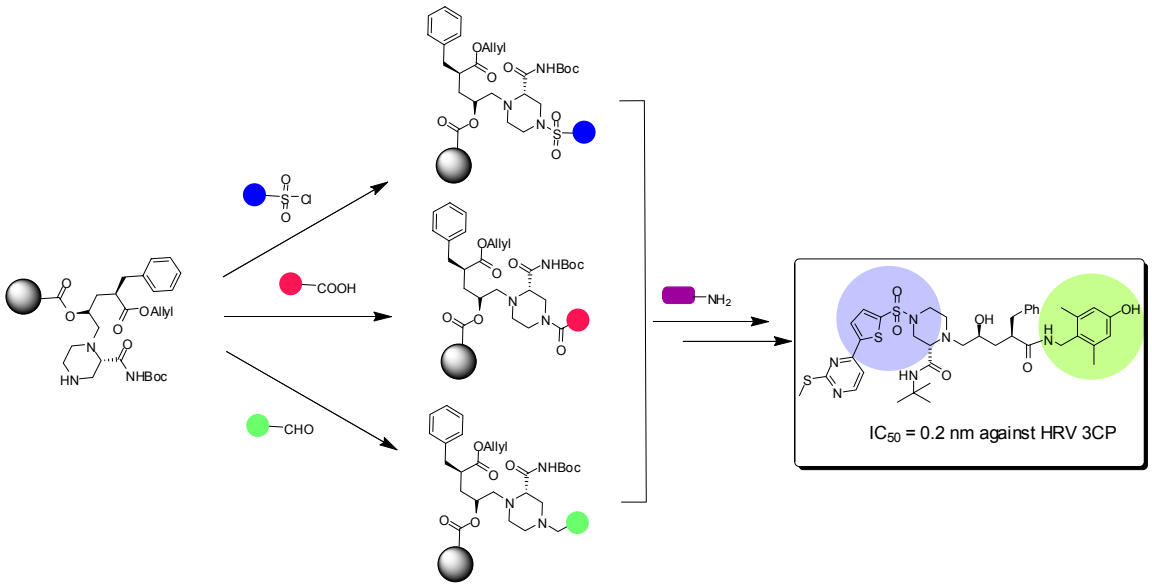


Fig. 3.2 Solid-phase synthesis of Indinavir® analogues

3.2.4 Design of our traceless solid-phase library

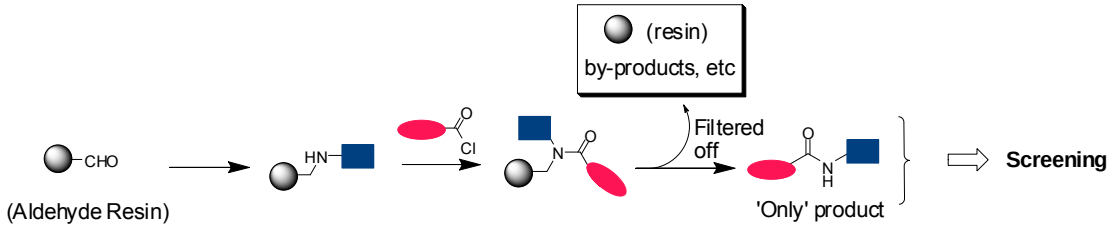


Fig. 3.3 Yao's solid-phase and *in situ* screening approach

In our strategy, we used the commercially available 4-Formyl-3-methoxyphenoxy (FMP) resin to capture various amine fragments via reductive amination, followed by attaching the isoxazole warhead. On cleaving the products form the resin, 'only' products

were obtained (**Fig 3.3**). Key advantages of our method include (i) it is a traceless approach allowing the use of exact same sets of starting material as in solution-phase synthesis; (ii) it's solid-phase, enabling a large library to be constructed efficiently; (iii) it's robust, giving high-quality products which in most cases are spectroscopically pure enough to be used directly for biological screening; (iv) 'only' products are taken for screening. The overall graphical representation of our strategy is shown in **Fig. 3.4**.

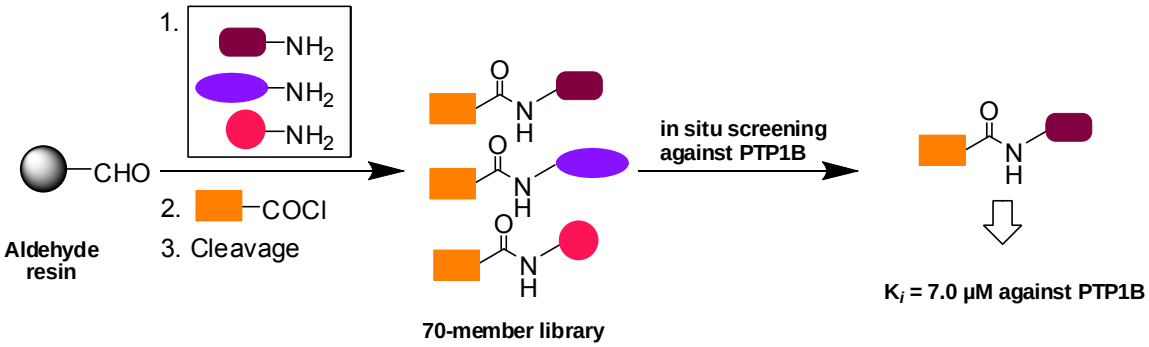


Fig 3.4 Graphical representation of our solid-phase strategy

3.2.4 PTP Inhibitor Design

In the current work, we have successfully developed our 2nd-generation, cell-permeable PTP1B inhibitors by using a solid-phase amide-forming reaction to rapidly link the core (acid functionalized isoxazole unit) and the peripheral groups (aromatic amine blocks) together, again here the inhibitors are bidentate one (**Fig. 3.5a**). Unlike traditional solution-phase approaches, our method delivers pure bidentate inhibitors at the end of synthesis and therefore is suitable for direct *in situ* screening. By screening these inhibitors against PTP1B, we have uncovered a candidate molecule which possesses comparable inhibition of K_i = 7.0 μM against PTP1B (**Fig. 3.5b**)[81].

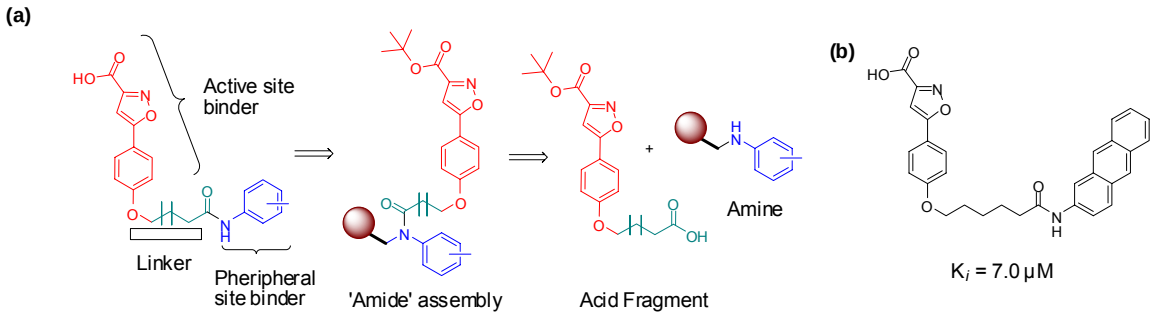


Fig. 3.5 (a) Inhibitor design- solid-phase amide bond formation reaction **(b)** Most potent inhibitor for the compound library

3.3 Results and Discussion

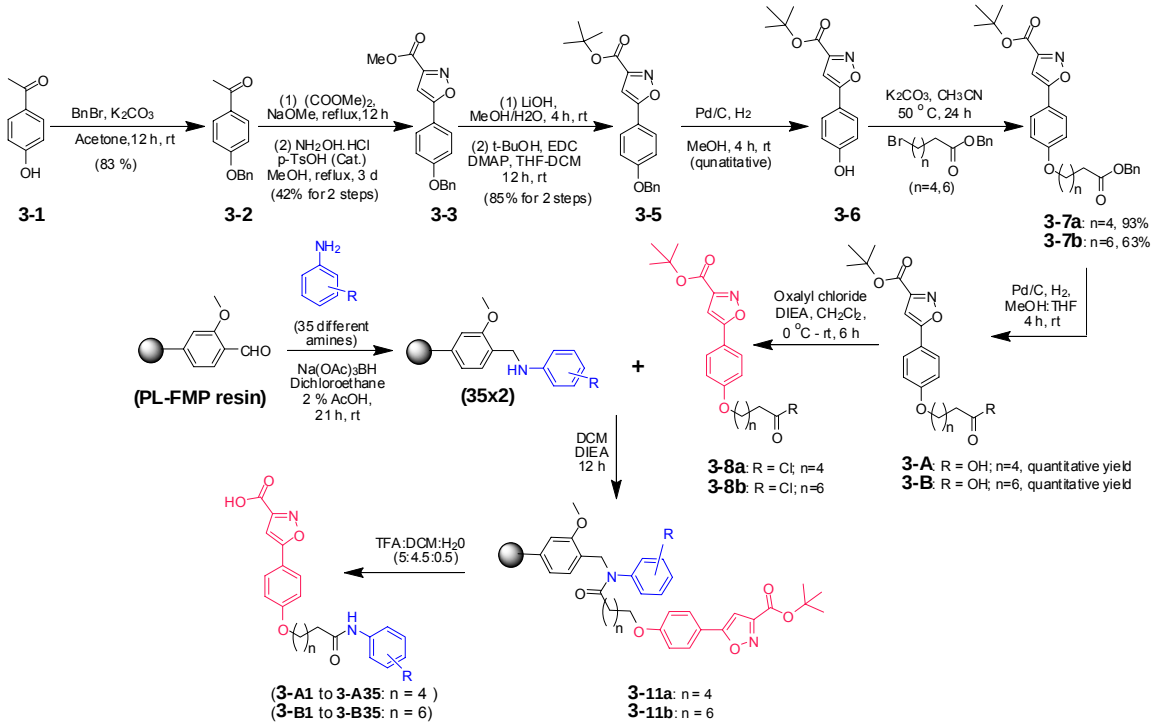
3.3.1 Chemical Synthesis of the inhibitor library

Synthesis of the acid functionalized war-heads

The synthesis of the acid-containing warheads, **3-A** and **3-B**, started from the commercially available 4-hydroxyacetophenone **3-1**, which on O-benzylation gave **3-2** (83% yield). Subsequently, condensation between **3-2** and dimethyloxalate in the presence of NaOMe, followed by cyclization with hydroxylamine gave the isoxazole carboxylic methyl ester, **3-3**, in modest yield (42% in two steps). Conversion of **3-3** to **3-5** was carried out first by base hydrolysis, followed by *t*-Bu ester formation (85% in two steps). Next, the benzyl ether **3-5** was deprotected in the presence of hydrogen (in Pd/C) giving **3-6**, followed by O-alkylation with two different linkers to afford **3-7a** and **3-7b** (93% and 63% yield, respectively). Subsequent deprotection of the benzyl esters gave the two acid-containing warheads, **3-A** and **3-B**, respectively.

Synthesis of the traceless inhibitor library

To start the assembly on solid phase, two sets of thirty-five aromatic amines were treated with FMP resin in the presence of $Na(OAc)_3BH$/2% glacial acetic acid in DCE to give the corresponding secondary aromatic amines **3-10**. Reductive amination proceeded smoothly with a variety of aromatic amines bearing different subsituents including -OH, -OR, -SR, -F, -Cl, $-OCF_3$, $-CO_2R$ and –R. Benzyl, napthyl and anthracenyl amines too underwent reductive amination smoothly. Subsequent amide bond-forming reaction between the resin-bound secondary aromatic amines and the acid warheads (**3-A** and **3-B**) was found to be highly challenging and required extensive optimizations. A variety of coupling reagents including HATU, PyBop, HBTU, EDC and DIC were attempted, but none gave the desired products in satisfactory yield and purity. Fortunately, by *in situ* conversion of **3-A** and **3-B** into their corresponding acid chlorides, **3-8a** and **3-8b**, with one equivalent of oxalyl chloride (with DIEA in DCM), we were able to successfully couple them, giving the resin-bound inhibitors **3-11** with excellent purity (> 95% in most cases). Finally, cleavage of the products from the resin using an optimized TFA cleavage cocktail (TFA:DCM:Water = 10:9:1) gave a total of 70 bidentate PTP inhibitors, **3-A1** to **3-A35** and **3-B1** to **3-B35** (**Scheme 3.1**). Most crude products were characterized by LC-MS/NMR and used for direct *in situ* screening for PTP1B inhibition.



Scheme 3.1 Traceless solid-phase synthesis of PTP1B inhibitors

3.3.2 Results and conclusion of the screening experiments

K_i measurements of selected inhibitors:

The inhibitory activity of the bidentate inhibitors was determined using a standard fluorescence microplate assay as explained in the previous chapter. First, an inhibitor fingerprint of the 70-member library against PTP1B was obtained, from which six potent hits (**3-A21**, **3-A26**, **3-A34**, **3-A35**, **3-B7** and **3-B25**) were identified (**Fig. 3.5**).

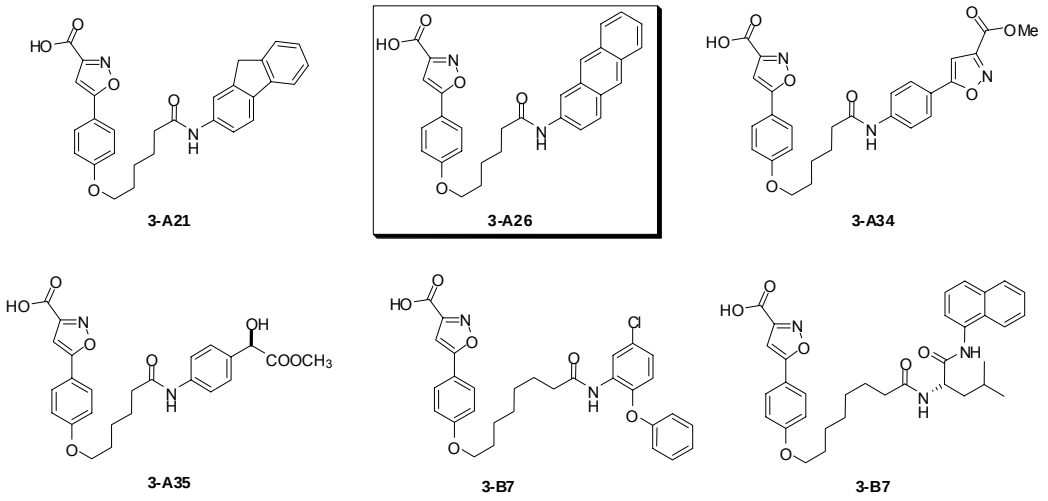


Fig. 3.5 Six candidate hits identified against PTP1B.

A kinetic evaluation was performed for three of the strongest inhibitors identified – **3-A21**, **3-A26** and **3-A34** against PTP1B. Dose-dependent reactions were carried out by varying both the concentration of substrate and inhibitor while keeping the enzyme concentration constant. Briefly, a two-fold dilution series from approximately 200 μM to 3.125 μM (final concentration) was prepared for each inhibitor. The enzymatic mix was incubated for 30 min before initiating the reaction with varying amount of substrate (2 μM - 8 μM). The enzymatic reactions were immediately monitored at λex = 355 nm and λem = 460 nm for a period of 5 min. The kinetic data obtained were analyzed using the Dixon Plot (1/initial velocity plotted against inhibitor concentration); affording three different linear graphs (corresponding to the different substrate concentration) that converged on the X-axis to give a value that corresponds to –Ki. The results are shown in **Fig. 3.6**.

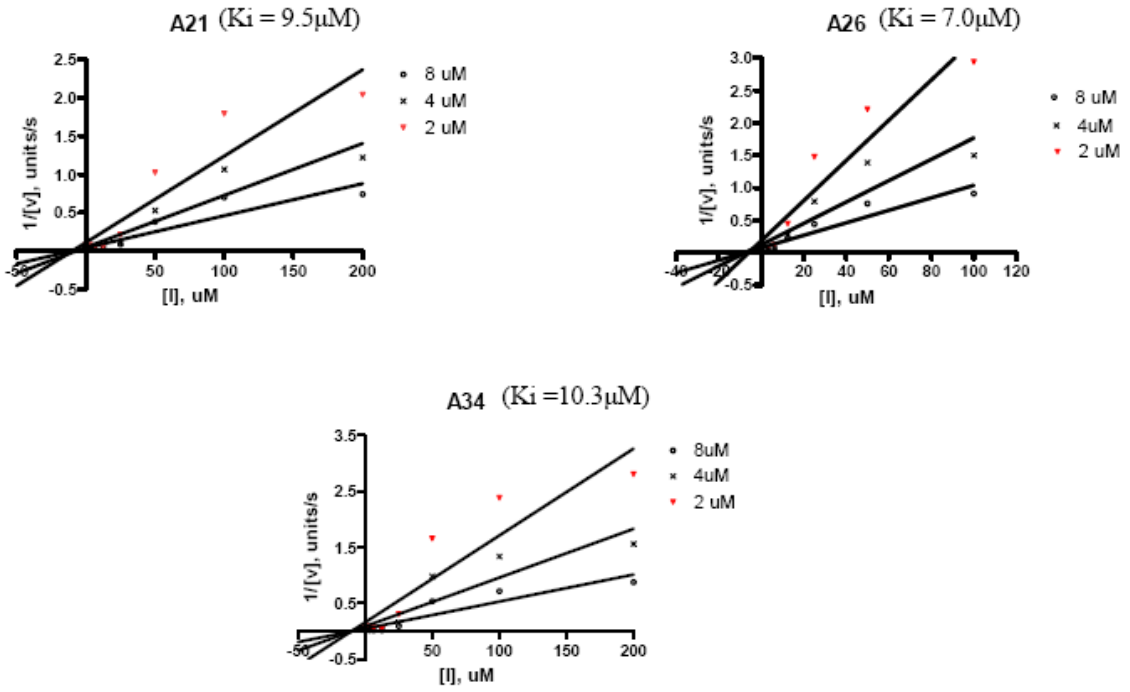


Fig. 3.6 Dixon plots for determination of K*i* values of three representative inhibitors against PTP1B

Discussion of the screening results

The best inhibitor against PTP1B was found to be **3-A26**, with IC_{50} and K*i* values of 10.3 μM and 7.0 μM, respectively (**Table 3.1**). Significantly, it also showed a 10-fold selectivity over TCPTP. It is interesting to note that **3-A26**, as well as the other good inhibitors of PTP1B, as shown in **Fig. 3.5**, contains a bulky aromatic group. This coincides reasonably with previously known cell-permeable inhibitors of PTP1B developed using other strategies, i.e. fragment-based or "click" chemistry approaches[15, 69, 70] thus further validating our solid-phase amide-forming methodology as a feasible strategy for future discovery of other enzyme inhibitors.

IC_{50} (K_i) in μM						
Enzyme	**3-A21**	**3-A26**	**3-A34**	**3-A35**	**3-B7**	**3-B25**
PTP1B	19.6 (9.5)	10.3 (7.0)	23.0 (10.3)	35.9	60.5	28.7
TCPTP	174.5	105.4	124.4	103.5	205.1	489.1

Table 3.1 Inhibition of the six selected inhibitors

3.4 Conclusion

In conclusion, we have successfully developed a solid-phase approach to assemble bidentate inhibitors of PTPs, by using amide-bond formation reaction. This strategy is highlighted by its robustness, product purity and allows diversification of a common core in a simple way. In our strategy, 35 different commercially available amines were captured via reductive amination on an aldehyde resin followed by coupling with two different acid-functionalized isoxazole warheads. Subsequent cleavage from solid-support provided a 70-member library which was taken directly for biological screening. Subsequent screening of the library against two PTPs (PTP1B and TCPTP) revealed three potent inhibitors with one compound (K_i = 7.0 μM) identified, whose inhibitoion against PTP1B was comparable to that of other known PTP1B inhibitors. Unlike the existing microtiter plate based approach, our method provides pure products which are free from the by-products, excessive starting materials and coupling reagents and avoids any false results during *in situ* screening. Also, our method uses the same sets of starting materials as in solution based microtiter plate synthesis. This solid-phase approach was developed

by us for the first time in the drug discovery area and applied to the successful discovery of PTP1B inhibitors. This method, which is superior to the existing methods, should find useful applications in high-throughput inhibitor discovery against various enzymes apart from PTPs.

Chapter 4

Versatile Microwave-Assisted Strategies for the Synthesis of Azide fragments

4.1 Summary

Libraries of small molecule fragments like azides, thiols, ureas, aryloxy/alkyloxy amines etc., are very crucial building blocks for the high-throughput fragment-based drug discovery. Only few synthetic methods/platforms are available to synthesize libraries of those molecules. We have devised simple and practical microwave-assisted strategies for the conversion of readily available alcohol building blocks into azides. In the first route, tosyl resin was used for the first time to synthesize azides in a catch and release approach. In the second route, alcohol building blocks were readily converted into azides via microwave (MW)-assisted azidation of tosylates/mesylates/chlorides. After a simple purification procedure, the azides synthesized from the above methods were suitable for directly click chemistry applications.

4.2 Introduction

Fragment-based drug discovery allows the modification or diversification of a core group (possessing moderate binding affinity towards a target protein/enzyme) using

simple fragments which possesses weak binding affinity to give a more potent bidentate type of inhibitors[62]. The small fragments could be either commercially and readily available building blocks like thiols, amines, aldehydes, acids, boronic acids etc or in some cases it could be molecules like azides (for click reaction), ureas (for Biginelli reaction), isocyanides (for Ugi multicomponent reaction), alkyl or aryloxyamines (for oxime formation reaction) which are difficult to access in large number for combinatorial library synthesis. Till-to-date no high-throughput methods are available for the versatile library synthesis of azides and ureas from simple and readily available starting materials.

4.2.1 Synthesis of azides – a literature review

Solid-phase azide synthesis

A combinatorial synthesis of 32 different aryl azides was accomplished by the $TMSN_3$ mediated cleavage of the polymer-bound aryl triazenes (**Fig. 4.1a**). The triazene can be modified in a number of ways, one such method is the Cu(I) mediated Ullmann–Nicolaou coupling[82]. These azides are convertible to a broad range of heterocycles. In another example, Suckling et al, synthesized heteroaryl azides from polymer supported heteroaryl sulfones in the presence of azide ions, by a nuclueophilic aromatic substitution reaction[83] (**Fig. 4.2b**).

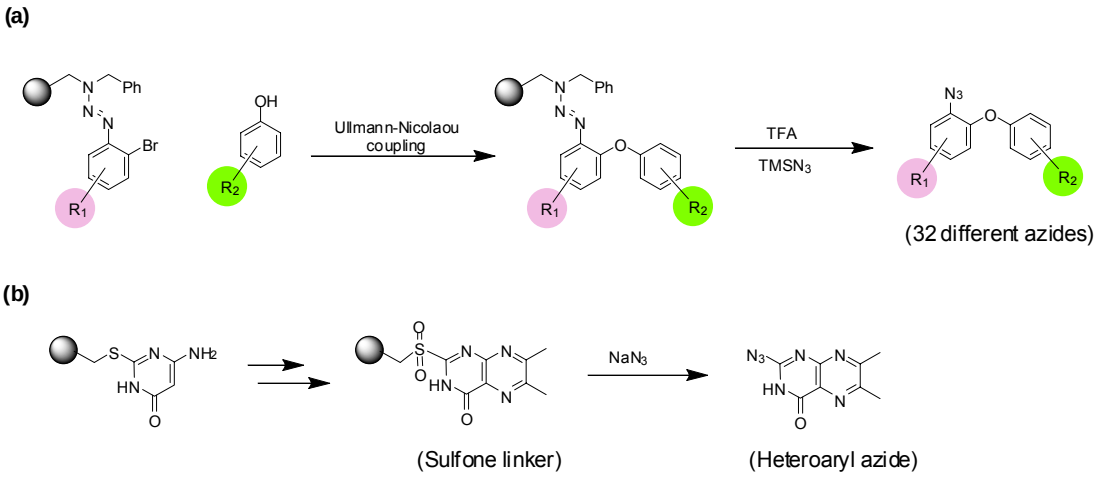


Fig 4.1 (a) Solid-phase synthesis of aryl azides from aryl triazenes **(b)** Solid-phase synthesis of heteroaryl azide from heteroaryl sulfone

Solution-phase and MW-assisted azide synthesis

In recent days microwave-assisted organic synthesis has a great impact on medicinal/combinatorial chemistry mainly because of its unique advantages like (1) Reduced reaction time (2) easy automation (3) energy efficiency[84]. Click reaction between an alkyne and azide in the presence of Cu (I) catalyst occupies a niche among the *in situ* inhibitor discovery, however as explained in the previous chapters the lack of efficient methods to synthesis libraries of azides in a high-throughput manner limits the application of click chemistry in the construction of inhibitor libraries. Apart from click application azides find a wide array of uses as versatile intermediates in organic synthesis[74], orthogonal amine protection[85] and Staudinger ligation[86]. Recently, improved methods for azide synthesis have shown some compatibility with click chemistry. For example Fokin *et al* reported the *in situ* generation of azides from aryl and alkyl halides followed by direct click chemistry[87, 88], however this method may not be suitable for *in situ* screening with enzymes since NaN_3 is a highly toxic chemical. Other methods include the direct conversion of amines to azides by employing the diazo-transfer

reaction using TfN_3[89, 90] or imidazole-1-sulfonyl azide hydrochloride[91], the use of TPP/DDQ/n-Bu_4NN_3 reagents[92] to convert alcohols directly to azides and the microwave (MW)-assisted azidation of tosylates[93, 94] (**Fig. 4.2**).

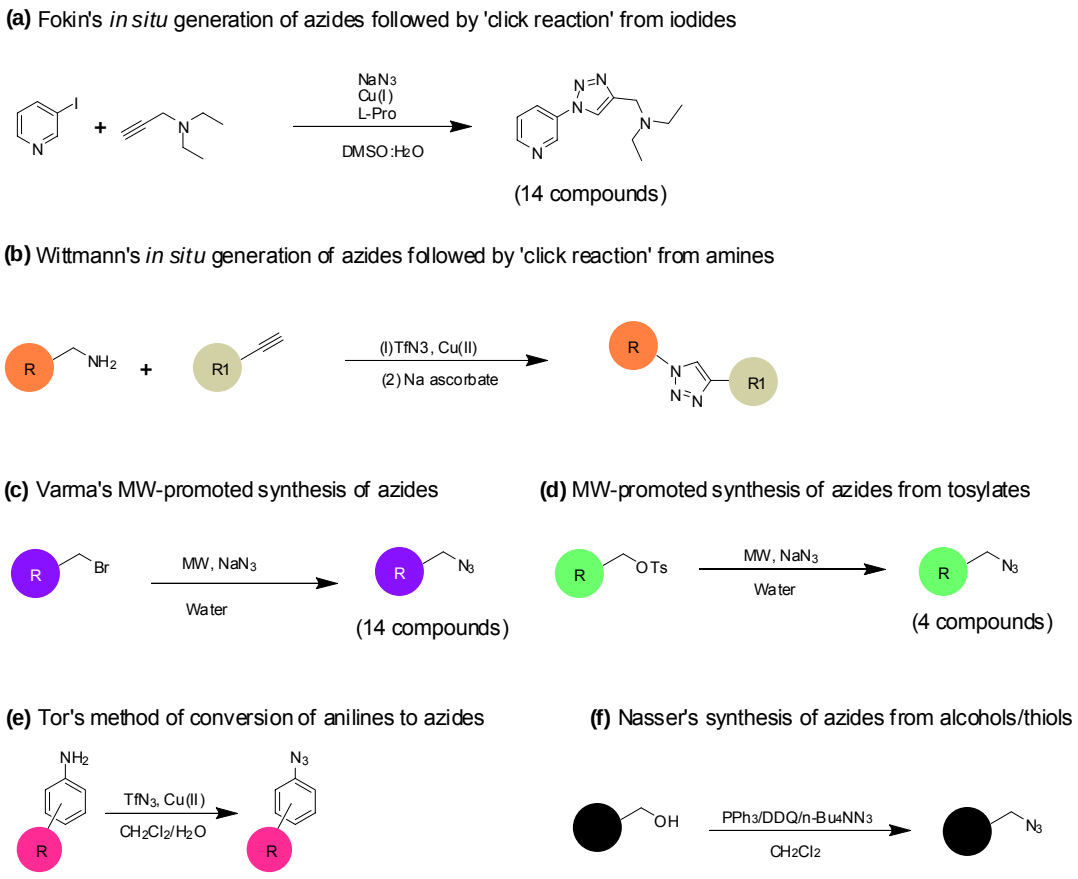


Fig 4.2 Various methods of synthesizing azides from commercially available building blocks

4.2.2 Drawbacks of the existing methods to synthesis azides

The above methods of generation of azides either use more expensive/unstable starting material/reagents (e.g. Tf_2O), or harsh conditions or needs column purification before or after click chemistry, which makes them unfit for rapid library generation. We have developed a MW-assisted traceless resin capture-release approach to generate azides and

also a practical and efficient MW-assisted approach to generate azides from readily available alcohol building blocks via tosylate/mesylate/chloride. We have also demonstrated the direct use of these azides in click reaction.

4.2.3 Design of our azide library

As shown in **Fig. 4.3a**, our first strategy made use of the commercially available PS-TsCl resin on which immobilized tosyl chloride moiety efficiently captured a variety of primary alcohols via simple nucleophilic substitution. The resulting resin-bound tosylates were subjected to a MW-assisted nucleophilic attack using sodium azide that released the alkyl azide from the resin. The alky azides were purified by simple extraction to remove the sodium azide and were directly used for click reaction. The advantage of this method includes (1) the method is traceless solid-phase strategy, allowing the use of as-is starting materials (i.e. same as solution-phase methods) as building blocks. (2) Variety of primary alcohols (except for some benzyl alcohols) can be converted to azides. (3) The MW-assisted azidation can give access to large library of azides in a short period of time. Our second strategy (**Fig. 4.3b**) is quite straight-forward, here primary and secondary alcohols were converted to tosylates/mesylates/chlorides using tosyl chloride or mesyl chloride and were purified by simple extraction and were subjected to rapid MW-assisted azidation in the presence of NaN_3 afforded the alkyl azides (primary and secondary) in very high purity after a simple extraction. The azides were directly used for the click reaction without any other purification (like column purification). The advantage of this method include (1) A variety of readily available 1^o and 2^o alcohols can be employed (2) MW- assisted azidation helps in the rapid synthesis of a library of compounds (3) a simple extraction is the only purification step.

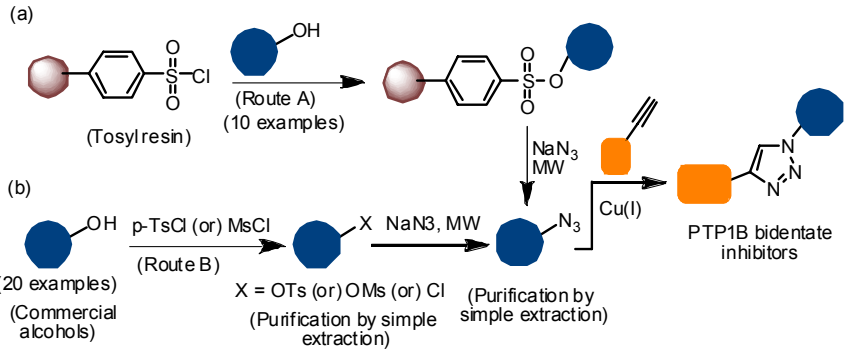


Fig. 4.3 MW-assisted synthesize of azides followed by direct click chemistry. **(a)** Solid-phase catch and release approach. **(b)** Solution-phase approach.

4.3 Results and Discussion

4.3.1 Chemical Synthesis of the azide library

4.3.1.1 Traceless solid-phase synthesis for PS-TsCl resin

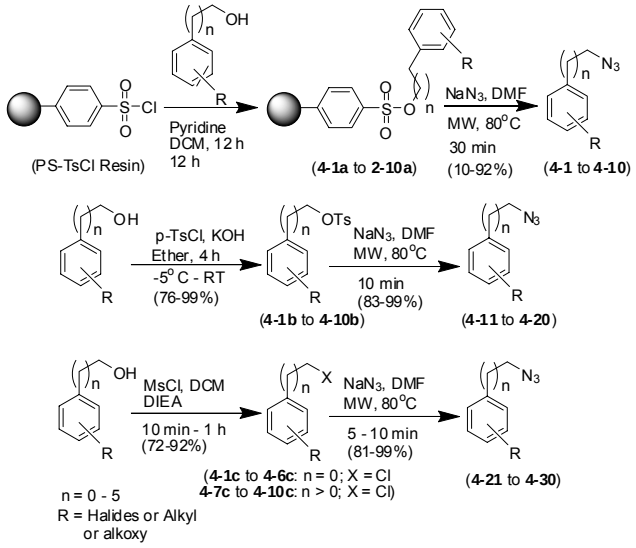
To demonstrate our traceless solid-phase strategy 10 different azides were synthesized from the various primary alcohols (**Scheme 4.1**). PS-TsCl Resin was treated with 6 equivalents of alcohol in the presence of 1:0.5 dichloromethane/pyridine to afford the tosylated resins (**4-1a** to **4-10a**). The tosylated products were subjected to MW- assisted azidation (80 ^{o}C for 30 min) in the presence of NaN_3 to release the alkyl azides which were purified by simple extraction to afford the product (**4-1** to **4-10**). However all the benzyl alcohol underwent poor conversion (0 to 20% yield) probably due to the steric hindrance between the aromatic moiety and the polymer backbone. Satisfactory yields (55-92%) and purity was obtained in the case of long chain alcohols (**Scheme 4.1a**). To demonstrate our second approach (**Scheme 4.1b**), 10 commercially available alcohol building blocks were treated with p-toluenesulfonyl chloride in the presence of KOH to

afford the tosylates (**4-1b** to **4-10b**) in good yield and purity after a simple extraction. The tosylates on MW assisted azidation (80 ^{o}C for 10 min) using sodium azide, followed by purification by simple extraction gave the azides (**4-1** to **4-4** and **4-11** to **4-16**) in good yield (83 to 99%) and purity.

4.3.1.2 Solution-phase MW-assisted azidation

10 different 1^{o} and 2^{o} alcohol building blocks were treated with mesyl chloride in the presence of DIEA. When benzyl alcohols are subjected to mesylation, surprisingly benzyl chlorides (**4-1c** to **4-6c**) were formed exclusively instead of benzyl mesylates, this would be because of the fact that benzyl mesylates are too reactive that they underwent a nucleophilic substitution with the chloride ion of the DIEA.HCl salt. Rest of the alcohols formed the mesylates (**4-7c** to **4-10c**) in high yield. The mesylates and the chlorides were accessible in less than an hour. Again purification of the mesylates/chlorides was done by simple extraction. It should be noted that no column purification were done to purify the tosylates, mesylates and the chlorides. The MW-assisted azidation (80 ^{o}C for 5-10 min) of chlorides/mesylates (**4-1c** to **4-10c**) using sodium azide followed by simple extraction gave the azides (**4-1**, **4-2**, **4-5** to **4-7**, **4-17** to **20**) in very good yield (81-99%) and purity. Three common azides synthesized through the different approaches were compared in **Table 8**. Short chain alcohol **4-a** under went conversion to azide **4-1** (via tosylate) in moderate yield when synthesized through method **A**. While good yields were obtained when the azide **4-1** was synthesized via method **B** and **C**. Long chain alcohols like **4-b** were converted to azide **4-2** in excellent yield and purity when synthesized via all the above three methods. Conversion of secondary alcohol **4-g** to secondary azide **4-7** was

unsuccessful in method **A.** but method **B** gave the secondary azide **4-7** in moderate yield and purity, while method **C** afforded the secondary azide **4-7** in excellent yield and purity (**Scheme 4.1c**). The structure of the azides synthesized by all the mentioned approaches and the yields and purity are given in **Table 4.1**



Scheme 4.1 (a) Solid-phase synthesis of azides from P-TSCl resin **(b)** Synthesis of azides via tosylates **(c)** Synthesis of azides via mesylates/chlorides.

4.3.1.3 Utilization of the azides in click assembly of PTP inhibitors

The feasibility of the above synthesized azides to be directly used for click chemistry application was proved by assembling PTP active site binding isoxazole functionalized alkynes (**2-F** and **2-A**) with the selected azides made from the above explained methods to get PTP1B bidentate inhibitors (**Scheme 4.2**). In all the three cases, the azides were completely consumed yielding the triazole products which are spectroscopically pure enough for biological screening experiments.

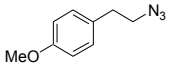
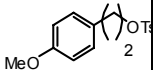
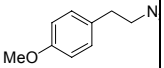
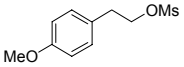
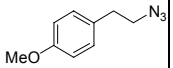
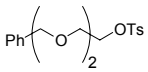
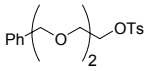
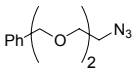
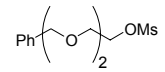
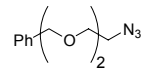
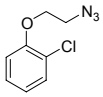
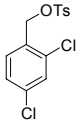
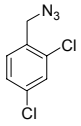
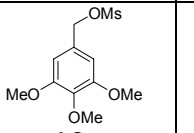
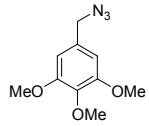
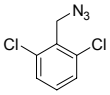
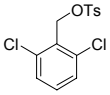
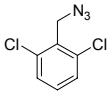
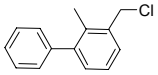
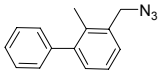
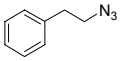
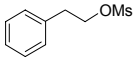
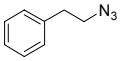
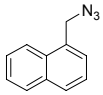
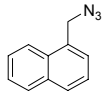
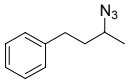
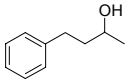
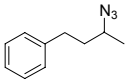
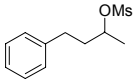
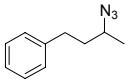
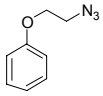
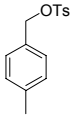
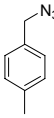
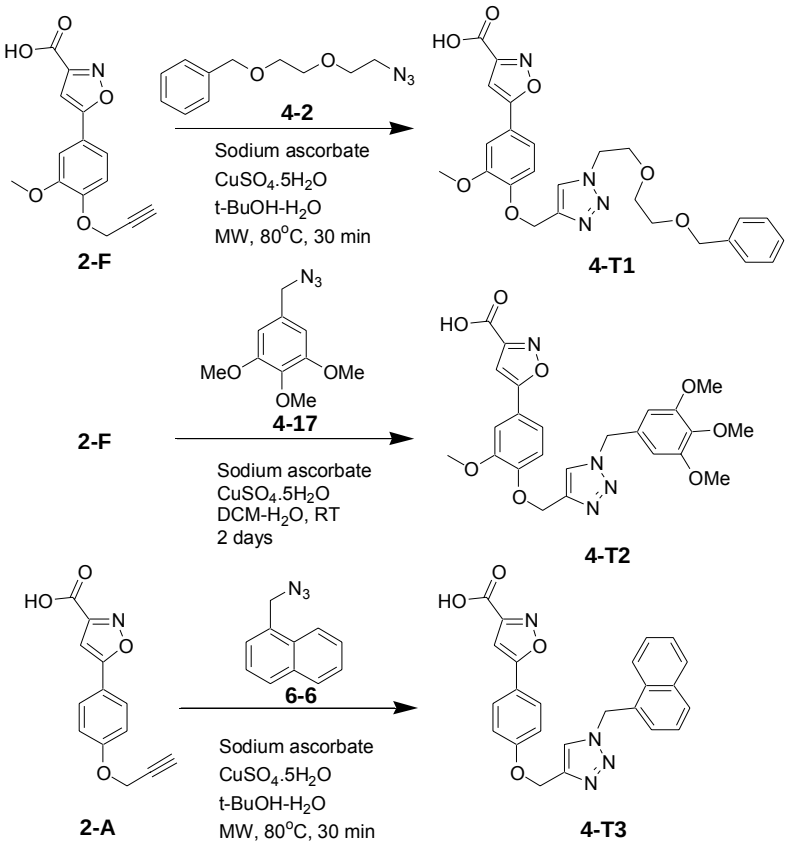
#	Method 1		Method 2a			Method 2b		
	Azide	Yield	Tosylate	Azide	Yield	Mesylate	Azide	Yield
1	 **4-1**	64%	 **4-1b**	 **4-1**	93%	 **4-1c**	 **4-1**	94%
2	 **4-2**	80%	 **4-2b**	 **4-2**	89%	 **4-2c**	 **4-2**	87%
3	 **4-3**	70%	 **4-3b**	 **4-11**	96%	 **4-3c**	 **4-17**	89%
4	 **4-4**	20%	 **4-4b**	 **4-4**	98%	 **4-4c**	 **4-18**	95%
5	 **4-5**	89%	 **4-5b**	 **4-12**	93%	 **4-5c**	 **4-5**	93%
6	 **4-6**	10%	 **4-6b**	 **4-13**	89%	 **4-6c**	 **4-6**	81%
7	 **4-7**	0%	 **4-7b**	 **4-7**	60%	 **4-7c**	 **4-7**	82%
8	 **4-8**	92%	 **4-8b**	 **4-14**	93%	 **4-4-8c**	 **4-19**	96%
9	**4-9**	61%	**4-9b**	**4-15**	96%	**4-9c**	**4-20**	92%
10	**4-10**	60%	**4-10b**	**4-16**	91%	**4-10c**	**4-21**	94%

Table 4.1 List of all azides and yields synthesized by three different approaches.



Scheme 4.2 Utilization of the azides in click assembly of PTP inhibitors

4.6 Conclusion

'Click' chemistry have conquered different areas of science, mainly medicinal chemistry, and relies on libraries of azide fragments for drug discovery applications. Till-to-date only very few methods are available for the rapid synthesis of azides, which calls for the development of new high-throughput methods to generate libraries of azides using simple building blocks. We have devised two simple methods to generate libraries of azide fragments using both solution-phase and solid-phase synthesis using readily available alcohol building blocks. In the first approach, alcohols were readily converted into azides by microwave-assisted azidation of mesylates/tosylates/chlorides and in the

second approach, for the first time, we have demonstrated the use of tosyl resin in the MW-assisted synthesis of azides by a simple catch-release approach. A total of 20 different azides were synthesized using the above approaches and their utility for direct click chemistry applications was demonstrated by the synthesis of PTP1B inhibitors. We believe that our new approaches would be a new high-throughput synthetic tool in the field of “Catalomics”.

Chapter 5

Activity-based fingerprinting of Enzymes

5.1 Summary

The first section of this chapter reports the design, synthesis of a library of 16 activity-based probes (ABPs) targeting different classes of proteases. These probes are designed for the first time to generate a reactive quinolimine methide upon hydrolysis of a peptide motif, which in turn covalently modify the enzyme. Further, these probes were used to generate unique substrate fingerprint profiles of proteases in gel-based proteomic experiments. The second section of this chapter narrates the synthesis of an activity-based tri-functional PTP probe based on *p*-hydroxy-mandelic acid unit and containing a biotin tag in addition to a fluorophore unit. This probe may find potential applications in identification as well as quantification of different PTPs.

5.2 Introduction

5.2.1 Activity-based fingerprinting

In the post-genomic era, characterization of enzyme activity pattern is more meaningful than enzyme identification. This is because the so-called substrate fingerprint of an enzyme reveals the type of chemical entities accepted by the enzyme as its potential

substrates, and thereby helps in a better understanding of its catalytic mechanism and properties. Similarly, the unique pattern generated for an unknown enzyme using a set of known substrates may be used to delineate its identity[95]. With the aid of standard analytical tools, traditional fingerprinting experiments use a whole spectrum of substrates and/or their analogs on a target enzyme and create quantitative and reproducible profiles directly related to the enzyme's activity. Different classes of enzymes have been studied in this fashion, including cytochrome P450, protein kinases, and hydrolytic enzymes[96, 97, 98]. In recent years, much effort has been expended in developing microarray-based bioassays[99]. When adopted for fingerprinting experiments, they could potentially provide a powerful platform by allowing the simultaneous analysis of thousands of enzymatic reactions on a single chip with very small sample volumes, while retaining a good degree of detection sensitivity[100].

Proteases, being one of the largest groups of enzymes which are important therapeutic targets of major human diseases, have been the focus of new enzyme assay developments in recent years[101]. Activity-based profiling (ABP), originally developed by Cravatt *et al* and adopted by other groups, allows studies of proteases present in a crude proteome on the basis of their enzymatic activities, rather than their relative abundance[102] ABP works by using either mechanism- or affinity-based chemical probes that enable their covalent attachment to different classes of enzymes, thus providing a versatile tool for large-scale protease identification, characterizations and even fingerprinting experiments[103]. **Fig. 5.1** shows the structures of some of the activity-based probes targeting various classes of proteases developed in Yao's and Cravatt's lab. Quinolimine derivatives activity-probes were used to study proteases in general[104]. Fluorophosphonate (FP) derivatives were used

as activity-based probes for detection of a broad spectrum of serine hydrolases[105]. Cysteine protease inhibitors, which include vinyl sulfone (VS)- and acrylamide-containing peptide conjugates were found to selectively target cysteine proteases[106]. Similar strategies were extended to the microarray format to allow potential high-throughput detection and identification of enzymes in a protein array[49, 99].

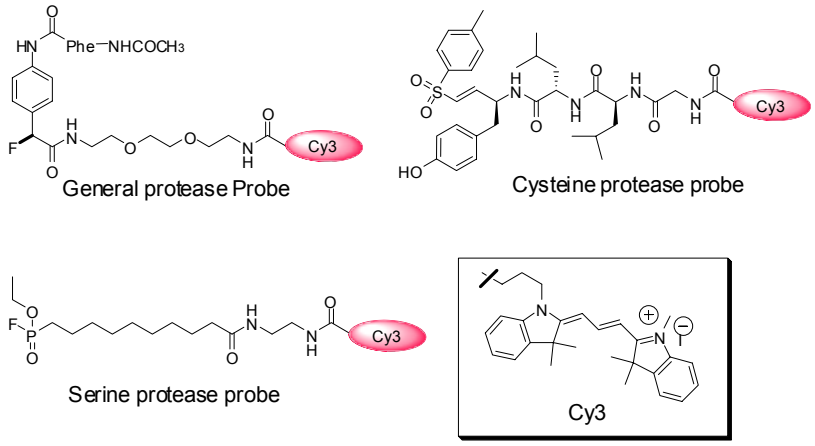


Fig. 5.1 Structure of the AB-protease probes developed by Yao's and Cravatt's lab

5.2.2 Design of the activity-based protease probe

We recently investigated a new class of ABP probes which target all major classes of proteases by their properties as enzyme substrates, rather than inhibitors. For this reason, we expect that these probes may be more suitable for protease substrate fingerprinting experiments than existing ones. This chapter illustrates our efforts in the first time the chemical synthesis of a full set of these probes, their utility for activity-based fingerprinting of proteases in gel-based experiments.

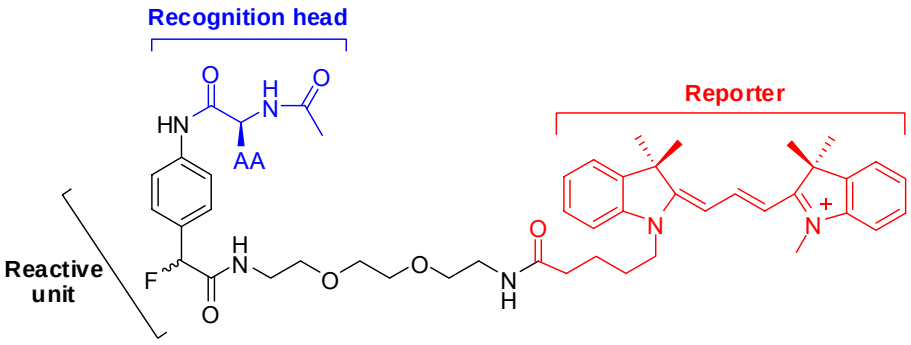


Fig. 5.2 Structure of the activity-based protease probe

Each of the probes contains a common *p*-aminomandelic acid moiety and a unique recognition head consisting of an N-acetylated amino acid that mimics the P1 position in a protease substrate. The amide bond between the two groups imitates the scissile bond in the protease substrate. A fluorescent reporter group, Cy3, was attached to the other end of p-aminomandelic acid. Upon proteolytic cleavage of the scissile bond, the probe releases the amino acid head group to generate a highly reactive quinolimine methide, which subsequently reacts covalently with the protease (that cleaves it) and renders it detectable (**Fig. 5.3**). One potential limitation of our enzyme-fingerprinting approach is that, since only P-site residues can be incorporated into the probe design, the approach might only be suitable for profiling proteases that possess P-site specificities.

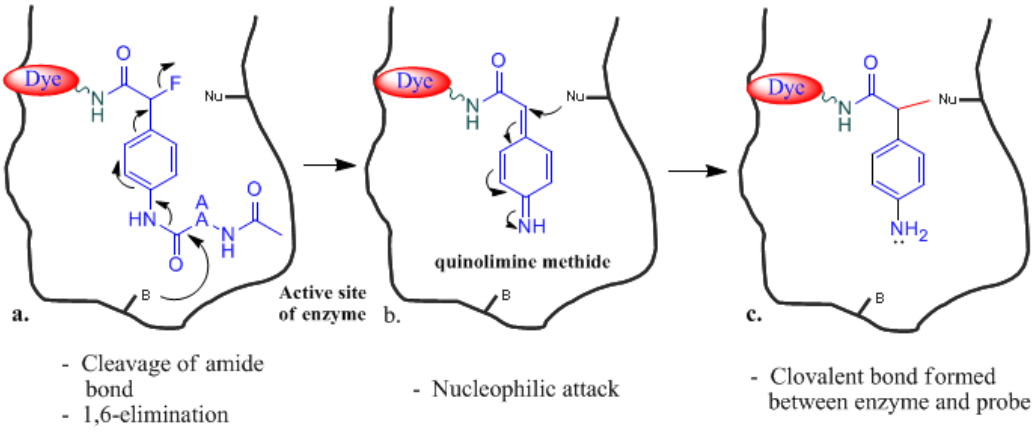
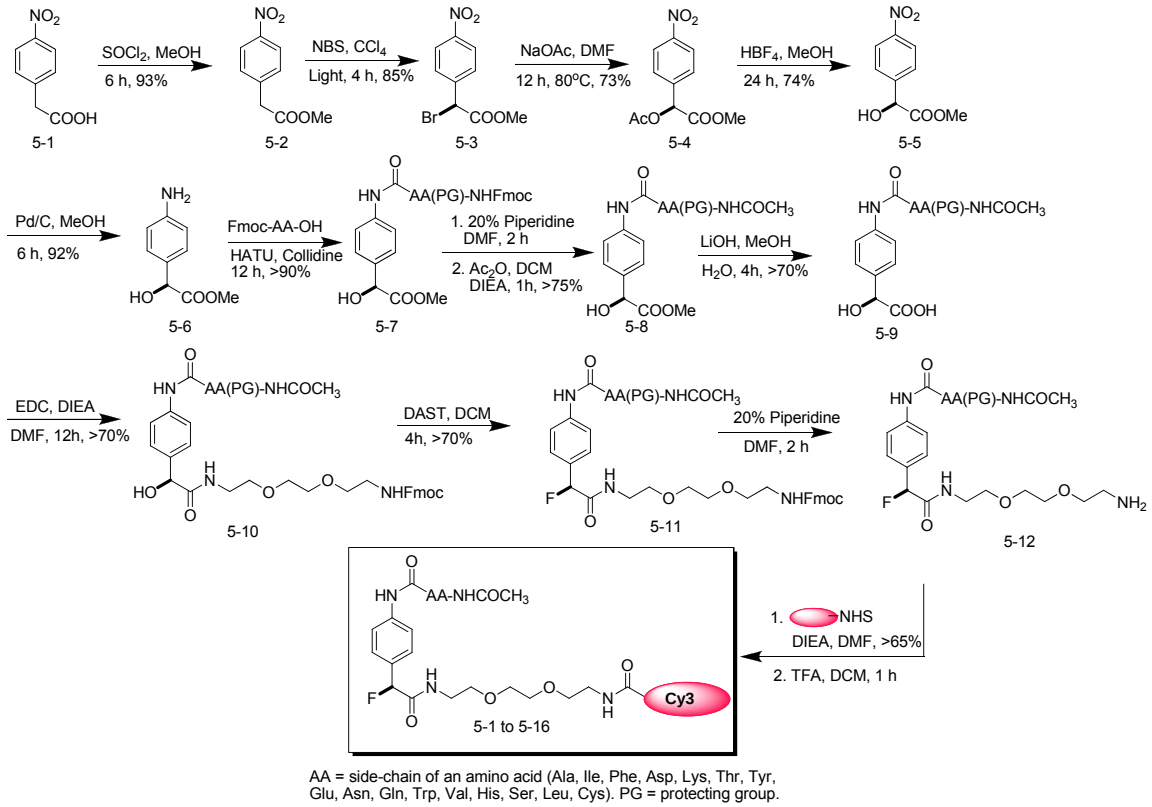


Fig. 5.3 Mechanism of labeling of the protease probe

5.3 Results and discussion

5.3.1 Chemical synthesis of the probes



Scheme 5.1 Synthesis of the activity-based protease probes

To make the sixteen probes, commercially available *p*-nitrophenylacetic acid, **5-17**, was treated with thionyl chloride in methanol to afford the methyl ester **5-18** in 93% yield. Subsequently, the benzylic proton in **5-18** was brominated by NBS (85% yield), generating **5-19**, followed by conversion to the corresponding benzyl alcohol **5-21** in two steps. Reduction of the nitro group in **5-21** with 10% Pd-C in the presence of H_2 gave the intermediate **5-22** with excellent yield (93%). A number of acylating reagents were tested in order to optimize the subsequent coupling reaction between the aromatic amine on **5-22** and a properly protected amino acid (both N- α -Boc and N- α -Fmoc amino acids were used), and it was found that HATU consistently gave the best yield (80-88% on average). Using the optimized method, a total of sixteen different amino acids were used to generate the same number of **5-23** in which each compound differs by its amino acid side chain. Upon deprotection of the N-α-Fmoc or N- α -Boc group, the resulting amino acids were acylated with acetic anhydride in DIEA to afford **5-24**. The hydrolytic cleavage of the methyl ester in **5-24** was achieved by using LiOH solution in nearly quantitative yield to furnish **5-25**. The final sixteen probes were subsequently obtained in three steps by conversion of the benzylic OH group in **5-25** to the corresponding fluoride with (diethylamino)sulfur trifluoride (DAST) at 0 ^{o}C, and attachment of a reporter Cy3 dye via a hydrophilic linker. The average yield of these three steps combined for all sixteen probes were approximately 50%. To ensure all final probes are free of impurities, they were subjected to preparative reverse-phase HPLC, followed by characterizations with ^{1}H-NMR and ESI-MS analysis as outlined in **Table 9**. In all, probes **5-1** to **5-16**, in which each contains a different N- α -acylated natural

amino acid as shown in **Scheme 5.1** (Ala, Ile, Phe, Asp, Lys, Thr, Tyr, Glu, Asn, Gln, Trp, Val, His, Ser, Leu, Cys), were successfully synthesized.

5.3.2 Fingerprinting Experiments

We next demonstrated this new class of substrate-based small molecule probes could be used to generate unique activity-based fingerprint profiles of different classes of proteases. Previously, only mechanism-based small molecule probes made of known protease inhibitors had been used in an ABP fingerprinting experiment. They were thus limited to studies of only certain classes of proteases. We expect that our current ABP probes, designed based on substrates rather than inhibitors of a protease, may be more suitable to generate fingerprints of the enzyme that better reflects its substrate specificity. These probes may also be used to fingerprint different classes of proteases simultaneously. We therefore labeled a total of thirteen different enzymes covering all four major sub-classes of proteases, each with our sixteen probes, **5-1** to **5-16**. The enzymes include serine proteases (i.e. α- & β-chymotrypsin, trypsin, thrombin, proteinase K and achromopeptidase), metalloproteases (i.e. actinase E and thermolysin), cysteine proteases (i.e. bromelain, papain and chymopapain), and aspartic proteases (i.e. renin and pepsin). With trypsin as a representative, a unique activity-based fingerprinting profile of the enzyme against all probes was thus generated. Based on the relative intensity of the labeled enzymes by each probe, a dendrogram, as shown in **Fig. 5.4**, were constructed using the programs Tree View™ and Cluster™ developed by Eisen and coworkers[107]. The thirteen proteases were hierarchically clustered based on the similarity of their fingerprint profiles, and shown by the tree structure on the left of the dendrogram. The

degree of similarity of protease was indicated as a function of height of the lines connecting profiles. In most cases, the labeling pattern of the enzyme by the probes correlates well with its previously known substrate preferences: α- and β-chymotrypsin, known to prefer hydrophobic/aromatic P_1 residues were labeled strongly by **F**, **Y** and **W**, as well as **Q**. The two enzymes were also expectedly clustered together by the software Cluster™, reaffirming their similarity. Bromelain, chymopapain, papain and proteianse K, by virtue of their known broad substrate specificities, were strongly labeled by a number of the probes and clustered together, once again indicating their similarity in substrate preference. Thermolysin, a metalloprotease known to tolerate a variety of p-site residues and instead prefer P'-site residues, was not clustered with any other proteases and surprisingly was labeled strongly by **E**. This underscores the potential drawback of our probes for fingerprinting certain classes of proteases. Trypsin, thrombin, achromopeptidase and actinase E all showed strong labeling by **K**, and were clustered together. This is in good agreement with these enzymes' known substrate preferences as well. Somewhat surprisingly, clustered in the same group were two other enzymes, pepsin and renin, which are known to have broad substrate specificities and require recognition elements on both P and P' sites. This again may indicate the limitation of our current class of probes. We previously showed that similarly designed probes, in which the enzyme recognition heads were made of peptides rather than single amino acids, confer much greater substrate specificity. In future, conceivably the same approach could be used to rectify problems faced here.

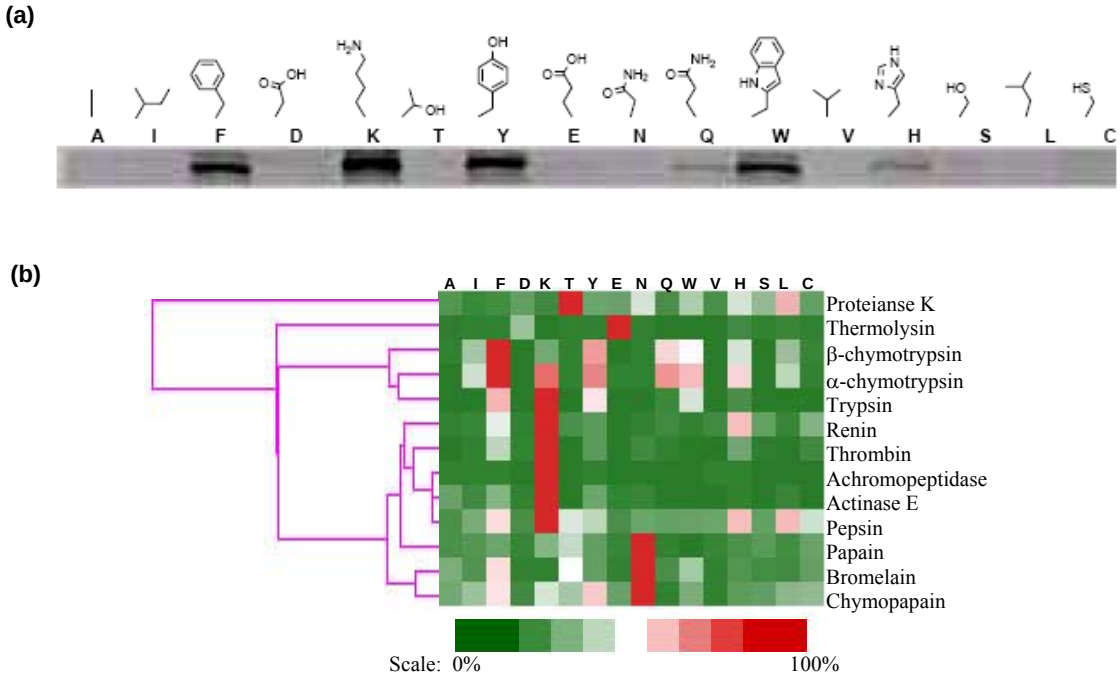


Fig. 5.4 (a) Affinity-Based labeling of Trypsin with probes 16 probes in the descending order. **(b)** Fingerprints of various proteases with probes **5-1** to **5-16**.

5.4 Activity-based Probes for Protein Tyrosine Phosphatases (PTPs)

Unlike proteases there are only very few reports on the activity-based probes for PTPs. 4-fluromethylaryl phosphate[108, 109] and α-Bromobenzylphosphonate[44] based ones are the two known AB-probes for PTPs (**Fig. 5.5a**). 4-fluromethylaryl phosphate based probes work by generating a highly reactive quinone methide intermediate on hydrolysis of the phosphate group by the 1,6-elimination of the fluoride. The quinone intermediate can potentially trap the phosphatases by the alkylation of the nucleophilic side chains near by the PTPs active site (**Fig. 5.5b**). The drawback of 4-fluromethylaryl phosphate probe is that it can form adduct with other phosphatases like Ser, Thr phosphatase calcineurin[103, 104, 105] due to its unspecific nature. Another main draw back is the diffusion

of the unmasked quinone methide intermediate, which could potentially alkylate other proteins in the vicinity that carry nucleophilic residues on their surface.

α- Bromobenzylphosphonate based probes work by generating an S-P bond between the active site cysteine of PTP and the P atom on the probe. Cysteine initiates a direct nucleophilic on the phosphonate group to give a transient phosphorane like intermediate (**Fig. 5.5c**), which undergoes a ring-closing reaction to form an epoxides like species with the expulsion of the bromide. Subsequent ring opening produces the hydroxylbenzylphosphonate covalently attached to the active site cysteine.

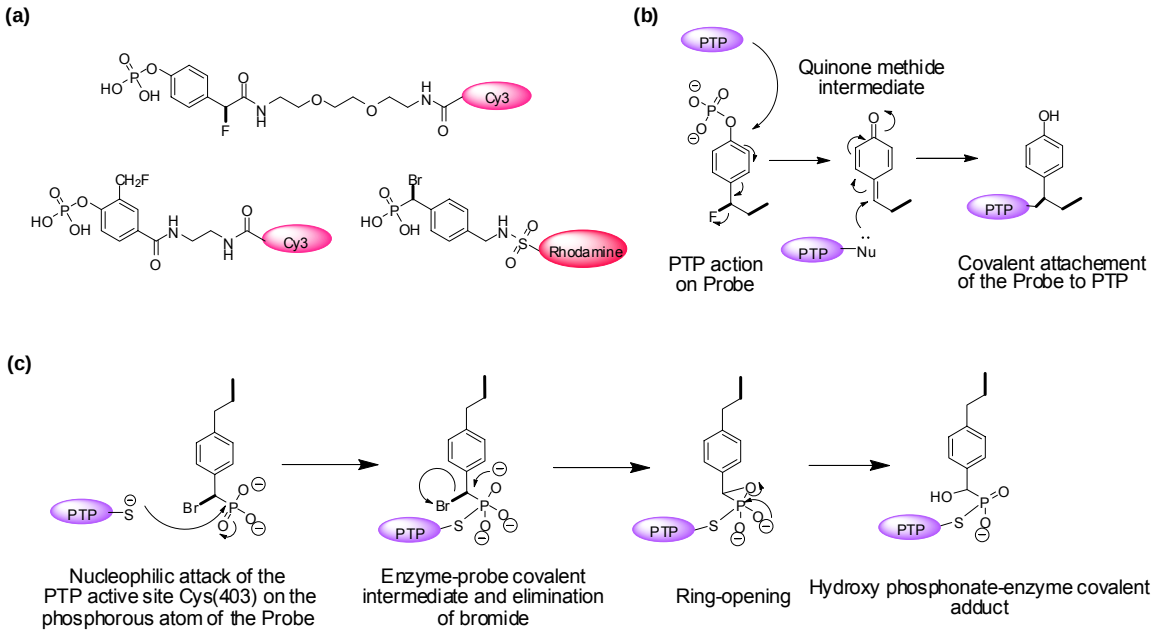


Fig. 5.5 (a) Structure of ABP PTP probes; Mechanism of the PTP labeling **(b)** fluromethylaryl phosphate probe **(c)** α-Bromobenzylphosphonate probe

5.4.1 Design of the PTP probes

The structure of the activity-based tri-functional probe consists of four units: a Cy3-containing fluorescence unit (for visualization purpose), biotin (for quantification purposes), a linker and a reactive unit (to capture PTPs). The fluorescence unit serves as a sensitive means to detect proteins upon labeling with the probe. The reactive unit is made of a 4-fluromethylaryl phosphate (mechanism-based group) that would covalently react with an enzyme in an activity-dependent fashion as shown in **Fig. 5.5b.**

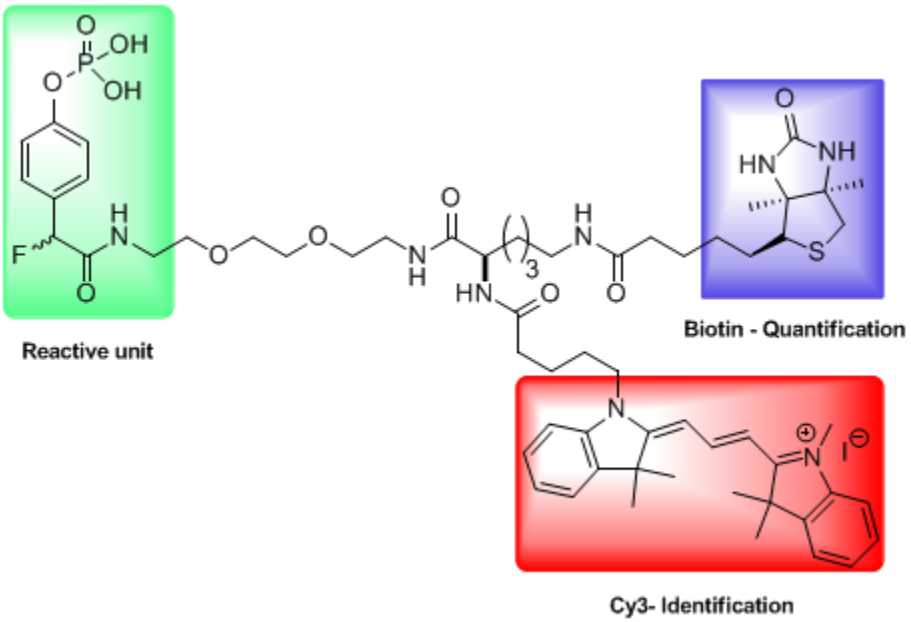
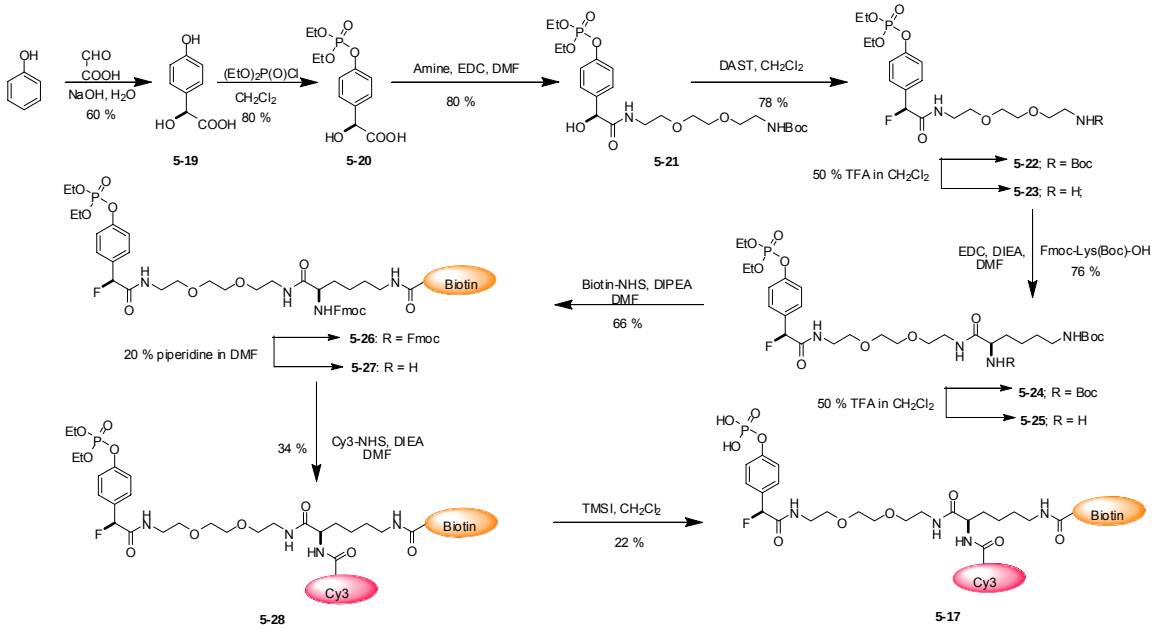


Fig. 5.2 Structure of the PTP tri-functional probe

5.4.2 Chemical Synthesis of the PTP probes

The tri-funtional probe **5-16** was prepared by adopting the below protocol (**Scheme 5.2**). Phenol on treatment with glyoxal under basic conditions gave the *p*-hydroxymandelic acid **5-19** which was phosphorylated using $(EtO)_2P(O)Cl$ to give **5-20**.

The acid 5-20 was coupled to an amine linker to give compound **5-21**, which was converted to Benzylic fluoride **5-22** by using DAST as a fluorinating reagent. Compound **5-22** on Boc deprotection furnished the TFA salt of the amine **5-23**, which was coupled to Fmoc-Lys(Boc)-OH using EDC as a coupling reagent to furnish compound **5-24**.



Scheme 5.2 Synthetic scheme for PTP tri-functional activity-based probe

Deprotection of the Boc protection on compound **5-24** offered compound **5-25** as a TFA salt, which was coupled with NHS ester of Biotin to give compound **5-26**. Compound **5-26** on treatment with 20 % piperidine in DMF furnished compound **5-27** which was coupled with the NHS ester of Cy3 to furnish compound **5-28**. The deprotection of the ethyl phosphonate ester of **5-28** was achieved using TMSI to yield the final probe **5-16**.

5.5 Conclusion

In conclusion, we have successfully synthesized and tested a panel of activity-based small-molecule probes that target different classes of proteases on the basis of their enzymatic activities and substrate specificities. These probes contain a peptide recognition motif that is linked via an amide bond to the *p*-aminomandelic acid moiety; enzymatic hydrolysis of the scissile peptide bond induces a facile rearrangement of the probe to generate the reactive quinolimine methide intermediates that in turn alkylates neighbouring residues within the enzyme active site resulting in the formation of covalent enzyme-probe adduct. These differ from previously developed ABP probes in that they were developed based on protease substrates rather than inhibitors; this makes them more suitable for generating activity-based fingerprinting profiles of proteases. In addition, they could be used on all different classes of proteases, a key advantage that is unattainable with other existing ABP probes. We have demonstrated that these probes are useful for generating unique substrate fingerprint profiles of proteases in gel-based proteomic experiments. Preliminary results indicate that they might be equally amenable for microarray-based enzyme-profiling experiments. We are currently investigating these probes' expanded utilities in proteomics research and attempting to address their potential limitations. The second part of the unit narrates the synthesis and of a tri-functional activity-based probe targeting PTPs, Briefly the structure of the probe consists of 4 units: a Cy3-containing fluorescence unit (for identification in gel-based experiments), a linker a reactive unit (based on *p*-hydroxymandelic acid) and a biotin tag (quantification). We believe that these probes may find more potential applications in high-throughput protein profiling as well as identification of unknown proteases and phosphatases.

Chapter 6

Bioimaging using small molecule probes

6.1 Summary

We have demonstrated the successful use of activity-based probes for the selective identification of protein classes in the previous chapter. Imaging of enzyme activities inside the live cells is very crucial for the understanding of their cellular mechanism and functional roles. This chapter demonstrates the design and synthesis of a total of seven small molecule probes designed for site-specific labeling of N-terminal cysteine-containing proteins expressed in live cells. Their utility for site-specific, covalent modifications of proteins was successfully demonstrated with purified proteins *in vitro*, and with live bacterial cells *in vivo* by taking advantage of the native chemical ligation.

6.2 Introduction

6.2.1 Native Chemical Ligation

Our own contribution to the field of bioimaging involves a novel strategy for site-specific covalent labeling of proteins *in vivo* [60, 110] by taking advantage of the chemoselective reaction of native chemical ligation[61]. The native chemical reaction is one of the very few non-enzymatic reactions that join two unprotected protein/peptide/small molecule fragments containing appropriately installed chemical functionalities, to

generate a ligated peptide/protein product with a native peptide bond at the reaction site. This highly chemoselective reaction occurs in an aqueous solution at physiological pH and involves a peptide fragment with an N-terminal cysteine residue and a second peptide fragment containing a C-terminal thioester group. The essence of the native chemical ligation reaction lies in the transthioesterification step between the thioester in one peptide and the sulfhydryl group from the N-terminal cysteine residue in the other to generate a ligated thioester intermediate, which undergoes spontaneous S→N acyl rearrangement to give rise to the final ligated product containing a native peptide bond at the ligation junction (**Fig. 6.1**). The transthioesterification reaction (i.e., the first step) is catalyzed by a suitable thiol additive (i.e., 2-mercaptoethanesulfonic acid), and is reversible under physiological conditions. The subsequent intramolecular nucleophilic attack by the a-amino group of the N-terminal cysteine to form the final amide bond is irreversible, and highly favorable due to the intramolecular five-membered ring formation and the subsequent generation of the thermodynamically stable amide bond. Consequently, all of the freely equilibrating thioester intermediates (e.g., from the first step) will eventually be depleted by the irreversible reaction in the second step, giving rise to only a single stable, ligated product. A key feature of this reaction is that it is highly chemoselective; the reaction occurs exclusively at the N-terminal cysteine of the peptide, even in the presence of other unprotected side-chain residues including internal cysteine residues. In our strategy, a protein of interest bearing an N-terminal cysteine is expressed inside a live cell by intein-mediated protein splicing [111] were selectively labeled by thioester containing small molecular probes. Our strategy provides an elegant

and simple way of site specifically labeling proteins in live cells, with little modification to the original protein apart from the addition of at most a few extra amino acids.

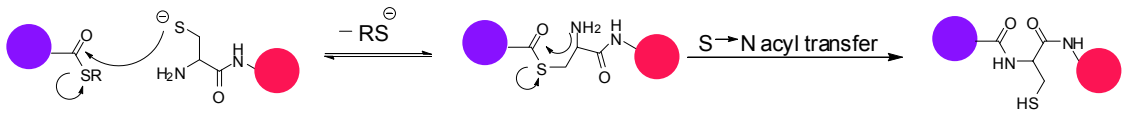


Fig 6.1 Native chemical ligation

6.2.2 Small-molecule design

In our strategy, a protein of interest having an N-terminal cysteine is expressed inside a live cell, by either intein-mediated protein splicing, or ubiquitin fusion[112]. Incubation of the cell with a thioester-containing, cell permeable molecule probe allows the probe to efficiently penetrate through the cell membrane into the cell, where the chemoselective native chemical reaction occurs predominantly between the thioester and the N-terminal cysteine of the protein, giving rise to the resulting labeled protein[110, 113]. Few endogenous N-terminal cysteine-containing proteins are known in various bacterial and mammalian genome databases, making our labeling strategy feasible in future for different live-cell imaging experiments[114]. Other endogenous molecules, such as cysteine and cystamine, are present in the cell and will also react with the probe However, their reaction products are also small molecules in nature, and can be easily removed, together with any excessive unreacted probe, by extensive washing of the cells after labeling.

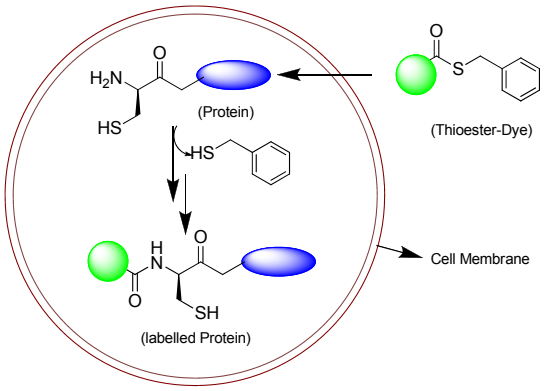


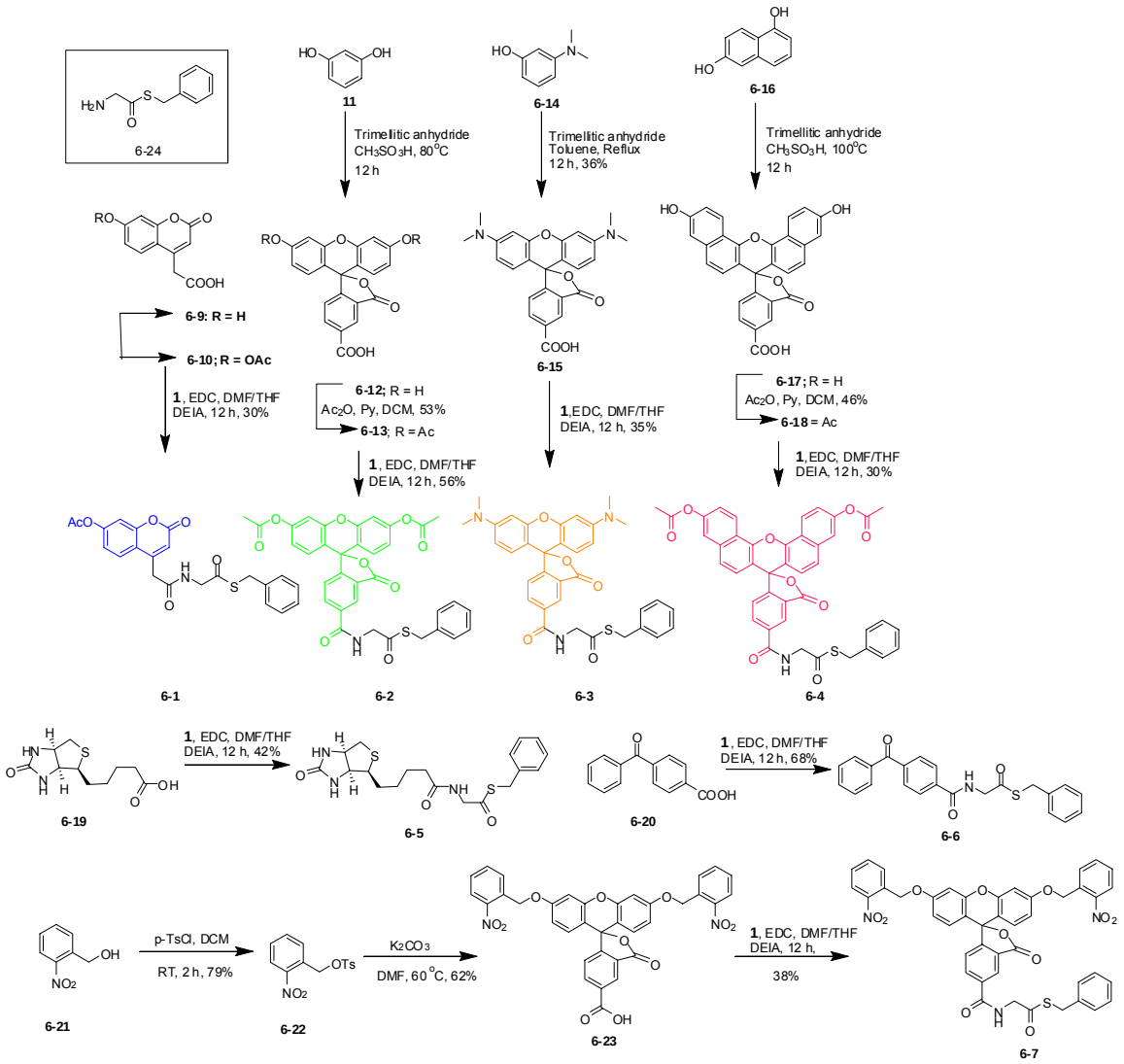
Fig. 6.2 Chemoselective reaction between a thioester-containing probe and an N-terminal Cys protein in a living cell.

6.3 Results and Discussion

6.3.1 Chemical Synthesis of the small molecule probes

A total of 7 different probes were (Scheme **14**), of which probes **6-1** to **6-4** are fluorophore-containing thioesters. **6-5** and **6-6** are biotin- and benzophenone-containing probes, respectively. **6-7** is a "caged" molecule of **6-2**, in which the fluorescence is designed to be "turned on" selectively upon photolysis. All probes were designed to be cell-permeable, in that acetates of different fluorophores were incorporated in **6-1, 6-2** and **6-4** to increase their cell permeability. The fluorophore in **6-3**, tetramethylrhodamine (TMR), as well as the biotin and benzophenone moieties in **6-5** and **6-6**, respectively, were previously shown to be cellpermeable[115]. Addition of the hydrophobic, benzyl-based thioester in all probes should further increase their cell permeability. Probes **6-1** to **6-4**, containing different fluorophores (*e.g.* coumarin (CM), fluorescein (FL), TMR and carboxynaphthofluorescein (CF), respectively) that emit in different colors, were designed for potential multicolor cell labeling and imaging. Proteins labeled with probes

6-5 and **6-6** may be used to study protein–protein interactions by *in vivo* experiments utilizing biotin–avidin binding and protein crosslinking, respectively. Probe **6-7** may be used for protein labeling in a live cell where temporal and/or confined fluorescence activation is needed[116]



Scheme 6.1 Synthesis of the cell-permeable probes for live cell labeling

6.3.2 Imaging experiments

In vitro labeling of proteins expressing an N-terminal cysteine was then carried out with the probes. A model protein, EGFP (enhanced green fluorescent protein), engineered

to contain an N-terminal cysteine, was incubated with probes **6-2, 6-3, 6-4** and **6-5** individually, and the extent of protein labeling was monitored over 24 hours by SDS-PAGE and Western blotting. 8 mM of each probe, with or without 1 mM of DTT, was added to the pure protein in 1 X PBS buffer. The reaction was quenched at specified time intervals with 10 mM of cysteine. Following protein separation on a 12% SDS-PAGE gel, the labeled protein was visualized and quantitated, either by a fluorescence gel scanner (in the cases of probes **6-2, 6-3** and **6-4**) or Western blotting using anti-biotin HRP conjugate and Amersham's ECL kit (in the case of probe **6-5**). Results are summarized in **Fig. 60** and **Fig. 61**. In all cases, the labeling was shown to reach near completion (> 75% labeling) within the first 3 hours of the reaction. In addition, more than 50% labeling occurred within the first 30 min of the reaction, making this strategy suitable for potential real-time bioimaging experiments in live cells. The site-specific nature of the labeling reaction was confirmed by repeating the experiment under identical conditions with control proteins which either do not have cysteine residues at all, or have only internal cysteines.

In all cases, labeling occurred ONLY with proteins possessing an N-terminal cysteine thereby unambiguously supporting our design principle, in which exclusive labeling should only occur at the N-terminal cysteine of the target protein. We next applied the strategy to the labeling of N-terminal cysteine proteins expressed inside live bacterial cells. *E. coli* cells overexpressing an N-terminal cysteine GST (glutathione-*S*-transferase) protein were treated with probe **6-3**, and the site-specific labeling of the probe inside the cells was assessed by fluorescence microscopy and SDS-PAGE (Fig. 2c and 2d, respectively): exclusive labeling occurred only with *E.coli* cells overexpressing N-

terminal cysteine GST. Minimal background labeling was unambiguously confirmed by SDSPAGE protein analysis of the labeled cells, indicating that the labeling occurred predominantly with the N-terminal cysteine GST.

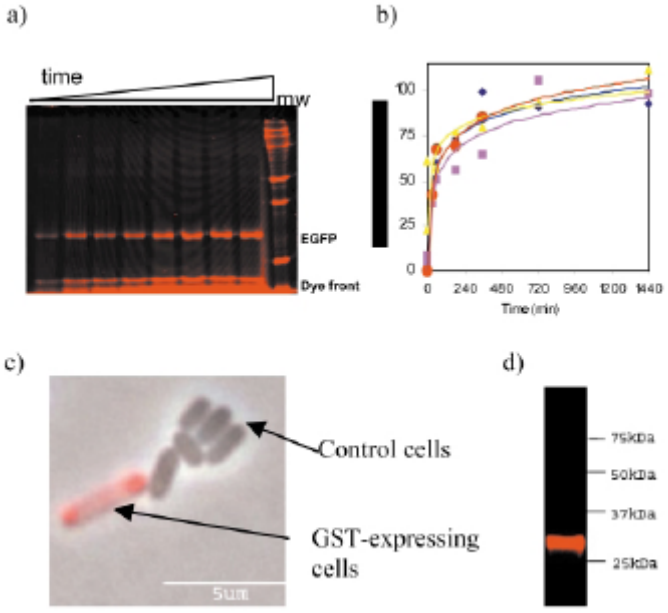


Fig 6.4 Site-specific labeling of N-terminal cysteine proteins with the small molecule probes. **(a)** SDS-PAGE of EGFP labeling with probe 4 for 1 min, 10 min, 30 min, 1 h, 3 h, 12 h, 24 h (left to right). **(b)** % completion of EGFP labeling over 24 h intervals with probe **6-2** (blue diamonds), **6-3** (purple squares), **6-4** (yellow triangles) and **6-5** (red diamonds). **(c)** Fluorescence microscope images of E. coli expressing N-terminal cysteine GST after labeling with probe 3 for 24 h. The fluorescence image was overlapped with the phase contract image for easy visualization. **(d)** 12% SDS-PAGE protein analysis of cells from **(c)**

6.4 Conclusion

Small molecule-based chemical labeling strategies for the site-specific tagging of proteins have provided powerful tools for the study of protein functions in intact cells. We have developed a simple yet highly versatile method for site-specific, covalent labeling of proteins inside live cells by adopting the chemo-selective native chemical ligation. We have shown that this strategy can be applied to bacterial as well as

mammalian cells with tolerable background labeling using a variety of small molecule probes, thus making it potentially useful for future bioimaging and proteomics applications. This strategy however is not amenable to proteins that have a pre-requisite for an amino-terminal signal sequence for organellar localization.

Chapter 7

Experimental Section

7.1 General Information

All chemicals were purchased from commercial vendors and used without further purification, unless otherwise noted. Either HPLC grade or distilled solvents were used in all the reaction. All moisture-sensitive reactions were performed under a positive pressure of nitrogen or argon. Concentration *in vacuo* was performed on a Büchi rotary evaporator Analytical thin layer chromatography was performed using Merck silica gel plates (0.25 mm thickness) with fluorescent indicator UV_{254}. Subsequent to elution, spots were visualized by ultraviolet illumination, iodine staining, $KMnO_4$ staining (for alkenes, aldehydes), Nindhydrin staining (for primary amines), or ceric molybdate staining. Flash chromatography was performed using Merck silica gel (40 μM particle size) and freshly distilled or AR grade solvents. Analytical and preparative RP-HPLC separations were performed on Phenomex C_{18} column (250 x 4.60 mm) and Phenomex C_{18} (250 x 21.20 mm) coloumns, respectively, using a Waters 600E HPLC system equipped with a waters 600 controller and a waters 2487 UV detector. Shimadzu LCMS-IT-TOF system equipped with auto-sampler (cat. No. LCMS-2010EV) was used for the analysis of the inhibitor library. Eluents A (0.1% TFA/acetonitrile) and B (0.1% TFA/water) were used as the mobile phase. ^{1}H NMR and ^{13}C NMR spectra were recorded on a 300 MHz

Bruker ACF300 or Bruker *Avance* 300 MHz or DPX-300 NMR spectrometer. Chemical shifts are reported as δ in parts per million (ppm) referenced with respect to residual solvent ($CHCl_3$ = 7.26 ppm, methanol-d4 = 3.3 ppm and DMSO-d6 = 2.50 ppm) or from internal standard tetramethylsilane (TMS = 0.00 ppm). ^{13}C NMR spectra are reported as δ in units of parts per million (ppm) relative from solvent signal: $CDCl_3$ (δ=77.0, triplet) and methanol-d4 (δ=49.0, singlet). The following abbreviations were used in reporting spectra: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of doublets. All enzymes were purchased from commercial sources.

7.2 High-Throughput Assembly of Protein Tyrosine Phosphatases (PTPs) Inhibitors Using "Click Chemistry"

7.2.1 Experimental details for the synthesis of PTP inhibitor library

General procedure for the synthesis of various Propargyl phenyl ethers (2-16, 2-19 and 2-22): Potassium carbonate (99 mmol) and benzo-18-crown-6 (3.3 mmol) were added to a solution of **propargyl** p-toluenesulfonate **2-34** (66 mmol) and hydroxyl acetophenone (66 mmol) in acetonitrile (100 ml) followed by refluxing for 6 h (*in the case of compound **2-21**, the reaction completed quantitatively in < 30 min), after which the organic phase was removed under reduced pressure and taken into dichloromethane layer (80 ml) and extracted with $NaHCO_3$ (2 x 40 ml), water (2 x 40 ml) and brine (1 x 40 ml). The organic phase was then dried with anhydrous Na_2SO_4, and the solvent was removed *in vacuo* to afford the crude product, which, upon further purification by flash

column chromatography, afforded the pure propargyl phenyl ether. Compounds **2-16**, **2-19** and **2-22** were prepared form compounds **2-15**, **2-18** and **2-21**, respectively, in 86–96% yield using above procedure.

1-(4-Prop-2-ynyloxy-phenyl)-ethanone (2-16): Yield = 89%. ^{1}H-NMR (300 MHz, CDCl3) δ 7.92 (d, *J* = 10.44Hz, 2H), 6.99 (d, *J* = 10.44Hz, 2H), 4.73 (d, *J* = 2.5Hz, 2H), 2.54-2.53 (m, 4H). ESI-MS: *m/z* $[M+H]^+$ = 175.0.

1-(5-Chloro-2-prop-2-ynyloxy-phenyl)-ethanone (2-19): Yield = 91%.

1-(5-Fluoro-2-prop-2-ynyloxy-phenyl)-ethanone (2-22): Yield = 98 %. ^{1}H-NMR (300 MHz, $CDCl_3$) 7.46-7.42 (dd, *J* = 3.15 & 3.12Hz, 1H), 7.19-7.12 (m, 1H), 7.06-7.01 (dd, *J* = 3.84 & 3.84Hz), 4.77 (d, *J* = 2.43Hz, 2H) 2.63 (s, 3H), 2.55 (m, 1H). ESI-MS: *m/z* $[M+H]^+$ = 193.2.

General procedure for the synthesis of various isoxazole-3-carboxylic acid methyl esters (2-17, 2-20, 2-23 and 2-25): To a mixture of alkyne-derivatized acetophenone (53 mmol) and dimethyl oxalate (53 mmol) was added freshly prepared NaOMe (0.5 M in MeOH, 53 mmol). The reaction mixture was refluxed for 12 h before cooled to room temperature. To the same reaction MeOH (80 ml), $NH_2OH.HCl$ (53 mmol) and a catalytic amount of *p*-$TsOH.H_2O$ were added and the resulting mixture was refluxed continuously for 3 days. Upon cooling to room temperature, the precipitated compound was collected, washed with water and ice-cold methanol to afford the pure isoxazole-3-

carboxylic acid methyl ester**.** Compounds **2-17**, **2-20**, **2-23** and **2-25** were obtained from compounds **2-16**, **2-19**, **2-22** and **2-24**, respectively, in 34–41% yield.

5-(4-Prop-2-ynyloxy-phenyl)-isoxazole-3-carboxylic acid methyl ester (2-17): Yield = 35%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.92 (d, *J* = 8.82Hz, 2H), 7.38 (s, 1H), 7.16 (d, *J* = 8.85Hz, 2H), 4.91 (d, *J* = 2.4Hz, 2H), 3.92 (s, 3H), 2.50 (m, 1H). ESI-MS: *m/z* $[M+H]^+$ = 258.1.

5-(5-Chloro-2-prop-2-ynyloxy-phenyl)-isoxazole-3-carboxylic acid methyl ester (2-20): Yield = 34%. ^{1}H-NMR (300 MHz, $CDCl_3$) 7.97 (d, *J* = 2.79Hz, 1H), 7.40-7.46 (dd, *J* = 2.43 & 2.79Hz, 1H), 7.21 (s, 1H), 7.07 (d, *J* = 9.06Hz, 1H), 4.85 (d, *J* = 2.10Hz, 2H), 3.99 (s, 3H), 2.59 (m, 1H). ESI-MS: *m/z* $[M+H]^+$ = 291.8.

5-(5-Fluoro-2-prop-2-ynyloxy-phenyl)-isoxazole-3-carboxylic acid methyl ester (2-23): Yield = 36 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.75-7.71 (dd, *J* = 2.79 & 2.79Hz, 1H), 7.25 (s, 1H), 7.19-7.08 (m, 2H), 4.85 (d, *J* = 2.43Hz, 2H), 4.01 (s, 3H), 2.57(m, 1H). ESI-MS: *m/z*: $[M+H]^+$ = 276.1.

5-(4-Nitro-phenyl)-isoxazole-3-carboxylic acid methyl ester (2-25): Yield = 36%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 8.39 (d, *J* = 9.24Hz, 2 H), 8.25 (d, *J* = 9.24 Hz, 2 H), 7.81(s, 1H), 3.32(s, 3H).

5-(5-Chloro-2-prop-2-ynyloxy-phenyl)-1H-pyrazole-3-carboxylic acid methyl ester (2-29): To a mixture of acetophenone **2-19** (1.1 g, 0.53 mmol) and dimethyl oxalate (0.87 g, 0.53 mmol) was added freshly prepared NaOMe (0.5 M in MeOH, 0.53 mmol). The reaction mixture was refluxed for 12 h and then cooled to room temperature. To the same reaction pot MeOH (20 ml), Hydrazinium sulfate (0.67 g, 0.53 mmol) and a catalytic amount of *p*-TsOH.H_2O were added and the resulting mixture was refluxed continuously for 3 days before cooled down to room temperature. The precipitated compound was collected, washed with water and ice-cold methanol to afford the pure pyrazole-3-carboxylic acid methyl ester **2-29** (0.63 g, 41% in 2 steps). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.71 (d, *J* = 2.43Hz, 1H), 7.28-7.32 (dd, *J* = 2.46 & 2.76Hz, 1H), 7.20 (s, 1H), 7.05 (d, *J* = 8.7 Hz, 1H), 4.85 (d, *J* = 2.10Hz, 2H), 3.96 (s, 3H), 2.62 (m, 1H). ESI-MS: *m/z* $[M+H]^+$ = 291.0

5-(4-Amino-phenyl)-isoxazole-3-carboxylic acid methyl ester (2-26): Compound **2-25** (3.5 g, 14.1 mmol) was dissolved in glacial acetic acid (100 ml) and Pd/C (0.35 g, 10% wt) was added. The reaction was stirred for 4 h under hydrogen atmosphere. At the end of reaction, Pd/C was filtered out and the solvent was removed completely under reduced pressure to yield pure amine **2-26** (2.8 g; 91%). ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.53 (d, *J* = 7.23Hz, 2H), 6.93 (s, 1H), 6.66 (d, *J* = 7.23Hz, 2H), 3.9 (s, 3H). ESI-MS: *m/z* $[M+H]^+$ = 219.1.

5-[4-(5,5-Dimethyl-4-oxo-hexanoylamino)-phenyl]-isoxazole-3-carboxylic acid methyl ester (2-27): To a solution of **2-26** (2 g, 9.2 mmol), **2-36** (1.6 g, 9.2 mmol) and

HATU (3.7 g, 10.1 mmol) in DMF (30 ml) was added 2,4,6-Collidine (1.3 ml, 10.1 mmol). The mixture was stirred for 3 h. The solvent was removed *in vacuo* and the resulting slurry was taken into ethyl acetate (100 ml) and extracted with $NaHCO_3$ (2 x 50 ml), water (2 x 50 ml) and Brine (1 x 50 ml), dried with anhydrous Na_2SO_4 and concentrated *in vacuo* to afford a pale brownish crude product, which on flash column purification afforded pure **2-27** (3 g; 91%). ESI-MS: *m/z* $[M+23]^+$ = 396.9 and *m/z* $[M+H]^+$ = 375.0.

5-[4-(3-Prop-2-ynylcarbamoyl-propionylamino)-phenyl]-isoxazole-3-carboxylic acid methyl ester (2-28): Compound **2-27** (2.0 g) was treated with 1:1 (v/v) ratio of 30 ml TFA/ dichloromethane mixture for 30 min to give the free acid. Upon removal of the solvent *in vacuo*, the acid was used without further purification. Next, the acid (1.35 g, 4.9 mmol) was dissolved in DMF, and EDC (1.04 g, 5.4 mmol), HOBt (0.75 g, 5.9 mmol) and DIEA (0.95 ml, 5.4 mmol) were added. The resulting mixture was stirred for 30 min before propargyl amine (0.38 ml, 5.4 mmol) was carefully added using a gas-tight syringe. The resulting mixture was allowed to stir in dark for 8 h. At the end, the solvent was removed *in vacuo* and the resulting residue was taken into ethyl acetate (80 ml) followed by extraction with $NaHCO_3$ (2 x 40 ml), water (2 x 40 ml) and brine (1 x 40 ml) and dried with anhydrous Na_2SO_4. Upon concentration *in vacuo*, the crude product was further purified by flash-column chromatography to furnish **2-28** (1.24 g; 74% yield). ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.90-7.87 (d, *J* = 8.70Hz, 2H), 7.76-7.73 (d, *J* = 8.73Hz, 2H), 7.36 (s, 1H), 3.92 (s, 3H), 3.34 (s, 2H), 3.56 (m, 4H), 2.49 (m, 1H). ESI-MS: *m/z* [M+Na] = 378.1.

General procedure for the hydrolysis of the methyl esters (2-A, 2-B, 2-C, 2-D & 2-E): The methyl ester (4 mmol) was suspended in methanol (10 ml) and a NaOH solution (10 ml; 10 M solution) was added slowly and the reaction was stirred for 3 h before the pH was adjusted to ~2 using HCl (2 N solution) under cold conditions. The resulting precipitate was collected, washed with cold water, dried *in vacuo* to furnish the desired acid. Compounds **2-A**, **2-B**, **2-C**, **2-D** & **2-E** were prepared starting from **2-17**, **2-20**, **2-23**, **2-28** and **2-29**, respectively.

Compound 2-A: Yield = 87%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.91-7.85 (d, *J* = 9.6Hz, 2H), 7.26 (s, 1H), 7.16-7.14 (d, *J* = 8.7 Hz, 2H), 4.90 (d, *J* = 1.98Hz, 2H), 2.50 (m, 1H). ESI $[M+1]^+$ = 244.1, ESI-MS: *m/z* $[M+Na]^+$ = 266.0.

Compound 2-B: Yield = 89%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.92-7.91 (d, *J* = 2.89Hz, 1H), 7.64-7.60 (dd, *J* = 2.79 & 2.76Hz, 1H), 7.36-7.33 (d, *J* = 9.03Hz, 1H), 7.18 (s, 1H), 5.09 (d, *J* = 2.1Hz, 2H), 2.51 (m, 1H). ESI-MS: *m/z* $[M+H]^+$ = 278.2.

Compound 2-C: Yield = 91%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.76-7.72 (dd, *J* = 2.76 & 3.15Hz, 1H), 7.74-7.40 (m, 1H), 7.34-7.31 (m, 1H), 7.19 (s, 1H), 5.06 (d, *J* = 2.46Hz, 2H), 2.50 (m, 1H). ESI-MS: *m/z* $[M+H]^+$ = 261.0.

Compound 2-D: Yield = 88 %. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.91-7.88 (d, *J* = 8.70Hz, 2H), 7.77-7.74 (d, *J* = 8.73Hz, 2H), 7.37 (s, 1H), 3.34 (s, 2H), 3.56 (m, 4H), 2.49 (m, 1H). ESI-MS: *m/z* $[M+H]^+$ = 341.9.

Compound 2-E: Yield = 88%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.92 (m, 1H), 7.42 (m, 1H), 7.25 (s, 2H), 4.99 (d, *J* = 2.45Hz, 2H), 2.51 (m, 1H). ESI-MS: *m/z* $[M-H]^{-1}$ = 275.0.

5-Bromo-pentanoyl chloride (2-31): 5-bromopentanoic acid (3 g, 17 mmol) was dissolved in distilled chloroform (100 ml) and redistilled thionyl chloride (2.81 ml, 34 mmol) was slowly added in a period of 10 min using a dropping funnel. The reaction mixture was refluxed for 12 h. Upon fractional distillation, 5-bromopentanonyl chloride was collected between 115^{o} and 119 ^{o}C. Repeated fractional distillation of the crude liquid afford very pure **2-31** (2.03 g; 60%) as a dull brownish liquid. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 3.42-3.38 (t, *J* = 12.21Hz, 2H), 2.96-2.92 (t, *J* = 13.23Hz, 2H), 1.94-1.74 (m, 4H).

General procedure for the preparation of azides (2-1 to 2-13): The amine (1 mmol) was dissolved in distilled chloroform (30 ml). The acid chloride linker (1.1 mmol) was slowly added at 0 ^{o}C, followed by addition of pyridine (1.1 mmol) to scavenge any protons generated in the reaction. The reaction mixture was allowed to stir for 1 h at room temperature. The organic layer was extracted with $NaHCO_3$ (2 x 20 ml), water (2 x 20 ml) and brine (1 x 20 ml), dried with anhydrous Na_2SO_4, then concentrated *in vacuo*. The resulting compound was pure enough to be used for the next reaction. The product obtained (1 mmol) was dissolved in DMF (30 ml). Sodium azide (1.1 mmol) was added and the reaction mixture was heated at 50-60 ^{o}C for 24 h. At the end, the solvent was

removed *in vacuo* and the residue obtained was taken into ethyl acetate (40 ml) or dichloromethane (40 ml) layer, followed by extraction with $NaHCO_3$ (2 x 20 ml), water (2 x 20 ml) and brine (1 x 20 ml). The resulting organic layer was treated with anhydrous Na_2SO_4 and concentrated *in vacuo* to yield the crude product which was further purified by flash column chromatography to afford the corresponding azide (60 to 80 % in two steps).

Azide 2-1: Yield = 64%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.53 (d, *J* = 7.62Hz, 2H), 7.35 (m, 2H), 7.17 (m, 1H), 4.18 (s, 2H).

Azide 2-2: Yield = 72%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.50 (d, *J* = 8.04Hz, 2H), 7.32 (m, 2H), 7.11 (m, 1H), 3.41 (m, 2H), 2.35 (m, 2H), 1.94 (m, 4H). ESI-MS: *m/z* $[M+H]^+$ = 219.1.

Azide 2-3: Yield = 76%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.95-7.92 (d, *J* = 8.73Hz, 2H), 7.76-7.73 (d, *J* = 8.73Hz, 2H), 3.83 (s, 3H), 3.66-3.64 (t, 2H), 2.71-2.66 (t, 2H). ESI-MS: *m/z* $[M+H]^+$ = 248.3.

Azide 2-4: Yield = 64%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.37 (d, *J* = 5.22Hz, 2H), 7.71 (d, *J* = 5.19Hz, 2H), 4.11 (s, 2H). ESI-MS: *m/z* $[M-H]^-$ = 176.1.

Azide 2-5: Yield = 63%.

Azide 2-6: Yield = 67%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.24 (q, 2H), 7.04-6.98 (m, 2H), 4.41 (d, *J* = 5.61Hz, 2H), 4.01 (s, 2H).

Azide 2-7: Yield = 77%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.55-7.52 (d, *J* = 8.55Hz, 2H), 7.29-7.26 (t, 2H), 7.09-7.05 (t, 1H), 3.62-3.58 (t, 2H), 2.59-2.55 (t, 2H). ESI-MS: *m/z* $[M+23]^+$ = 213.1

Azide 2-8: Yield = 82%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.57-7.54 (d, *J* = 8.43Hz, 2H), 7.34-7.31 (d, *J* = 8.73Hz, 2H), 5.06 (s, 1H), 4.04-3.98 (q, 2H), 3.62-3.58 (t, 2H), 2.68 (m, 2H), 1.14-1.10 (t, 3H).

Azide 2-9: Yield = 87%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.44-7.41 (d, *J* = 9Hz, 2H), 6.90-6.87 (d, *J* = 8.55Hz, 2H), 3.81 (s, 3H), 3.75-3.71 (t, 2H), 2.62-2.58 (t, 2H).

Azide 2-10: Yield = 93%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.41 (dd, *J* = 2.13 & 1.98Hz, 2H), 6.85 (dd, *J* = 2.28 & 1.98Hz, 2H), 3.80 (s, 3H), 3.31 (t, *J* = 13.32Hz, 2H), 2.37 (t, *J* = 14.64Hz, 2H), 1.61-1.85 (m, 4H). ESI-MS: *m/z* $[M+Na]^+$ = 271.0.

Azide 2-11: Yield = 66%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 8.01-7.97 (d, *J* = 11.85Hz, 2H), 7.50-7.21 (m, 6H), 5.03 (m, 1H), 3.87 (m, 2H), 2.74-2.24 (t, 2H).

Azide 2-12: Yield = 92%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.44 (d, *J* = 3.63Hz, 1H), 7.01 (d, *J* = 3.63Hz, 1H), 3.44 (t, *J* = 12.0Hz, 2H), 2.59 (t, *J* = 13.29Hz, 2H), 1.98-1.96 (m, 4H). ESI-MS: *m/z* $[M+H]^+$= 226.0.

Azide 2-13: Yield = 87%. ^{1}H-NMR (300 MHz, $CDCl_3$), δ 7.48-7.45 (d, 2H), 7.33-7.30 (d, 2H), 5.14 (s, 1H), 3.73 (s, 3H), 3.31-3.26 (t, 2H), 2.38-2.33 (t, 2H), 1.77 (m, 4H). ESI $[M+23]^+$ = 329.1.

Azide 2-14: Yield = 59%. ESI-MS: *m/z* $[M-H]^-$ = 393.0. The synthesis of this compound is as previously reported.[69]

General "click chemistry" procedure for library assembly

Each alkyne (13.5 μmol) and azide (16.2 μmol) were dissolved in DMSO (0.1 ml) and taken into *t*-butyl alcohol (1 ml) in a 15 ml centrifuge tube. Sodium ascorbate (1.35 μmol, dissolved in 1 ml of water, was added, followed by copper (II) sulfate pentahydrate (0.135 μmol). The resulting solution was shaken for 2 days. A few library members precipitated, which were centrifuged, washed with 1:1 cold methanol/water mixture and dried *in vacuo*. They were redissolved in DMSO and used directly in subsequent enzymatic assays. Unprecipitated reaction mixtures were also pure enough and used directly without further purifications. All reaction products were assessed by HPLC followed by MS, and shown to be of free of staring materials (e.g. the alkynes) and of high purity.

7.2.2 Biological assay for PTPs:

Screening for Inhibition activity against different phosphatases

The protein phosphatase activity was determined by measuring the rate of hydrolysis of the following fluorogenic substrate, designed and synthesized as previously reported[69] (**Fig. 7.1**).The inhibition of phosphatases were screened using black polypropylene flat-bottom 384-well microtiter plates (Nunc, USA) in a total reaction volume of 50 µl/well, monitored with a SpectraMAX^{TM} Gemini XS fluorescence plate reader (Molecular Devices, USA). Different concentrations of inhibitors were used in the assay (26-260 µM), where necessary. Controls were done with 1,3-dipolar ligation reactions carried out without addition of Cu(I) catalyst (e.g. $CuSO_4$ and sodium ascorbate), and showed no inhibition against any of the enzymes (up to 260 µM inhibitor concentration). The assay conditions were given below:

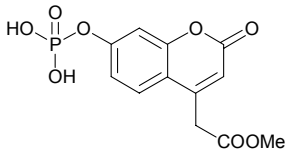


Fig. 7.1 Structure of the fluorogenic substrate used in the assay

PTB 1B:
PTB 1B concentration (1 U) = 1 µl
Substrate S1 (mM) = 1 µl
Inhibitor (**0.66** mM) = 2 µl
Assay Buffer (2 X) = 25 µl
Double Distilled water = 21 µl
[Assay buffer = 50 mM HEPES, 300mM NaCl and 0.2 mg/ml BSA, pH = 7.5).

TCP TP:
TCP TP concentration (0.33 U) = 1 µl
Substrate S1 (mM) = 1 µl
Inhibitor (**0.66** mM) = 2 µl
Assay Buffer (2 X) = 25 µl

Double distilled water = 21 µl
[Assay Buffer = 25 mM Tris, pH = 7.5]

YOP
YOP Concentration (1 U) = 1 µl
Substrate S1 (mM) = 1 µl
Inhibitor (**0.66-6.6 mM**) = 2 µl
Assay Buffer (2 X) = 25 µl
Double distilled water = 21 µl
[Assay Buffer = 25 mM Tris, pH = 7.5]

LAR
LAR Concentration (0.5 U) = 1 µl
Substrate S1 (mM) = 1 µl
Inhibitor (**0.66-6.6 mM**) = 2 µl
Assay Buffer (2 X) = 25 µl
Double distilled water = 21 µl
[Assay Buffer = 25 mM Tris, pH = 7.5]

λPPase:
λPPase Concentration (0.5 U) = 1 µl
Substrate S1 (mM) = 1 µl
Inhibitor (**0.66-6.6 mM**) = 2 µl
Assay Buffer (2 X) = 25 µl
Double distilled water = 21 µl
[Assay Buffer = 25 mM Tris, pH = 7.5]

PP1
PP1 Concentration (0.5 U) = 1 µl
Substrate S1 (mM) = 1 µl
Inhibitor (**0.66-6.6 mM**) = 2 µl
Assay Buffer (2 X) = 25 µl
Double distilled water = 21 µl
[Assay Buffer = 25 mM Tris, pH = 7.5]

The enzyme was last added to start the reaction and the reaction was monitored at λ_{ex} = 355 nm and λ_{em} = 460 nm. The activity was compared in the absence and presence of inhibitors. The results were shown in the heat maps below (**Fig. 7.2**):

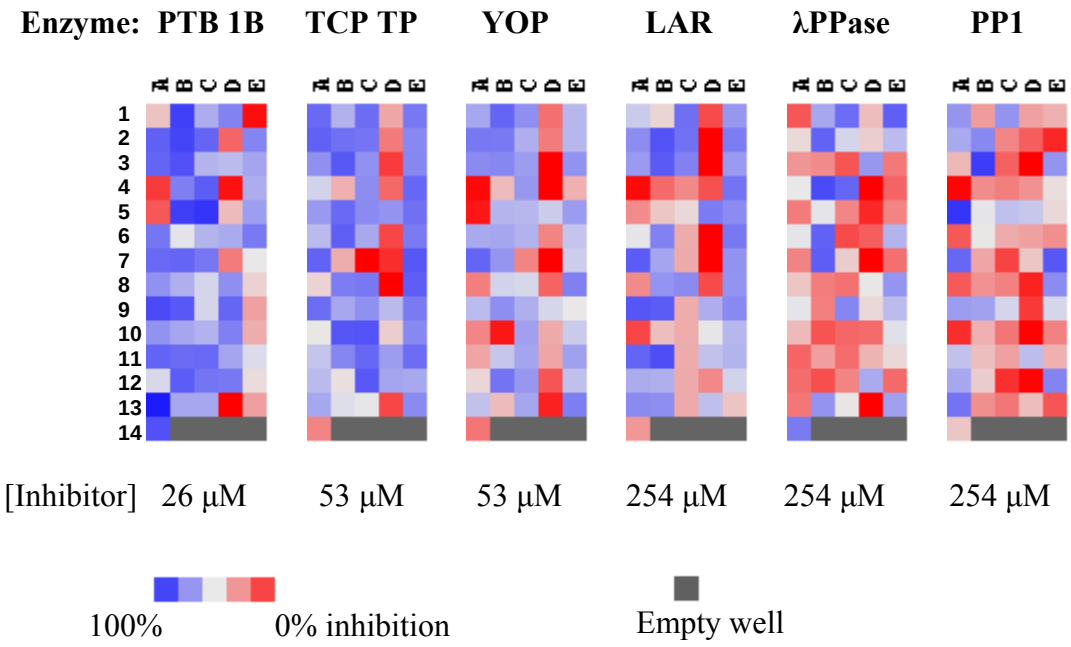
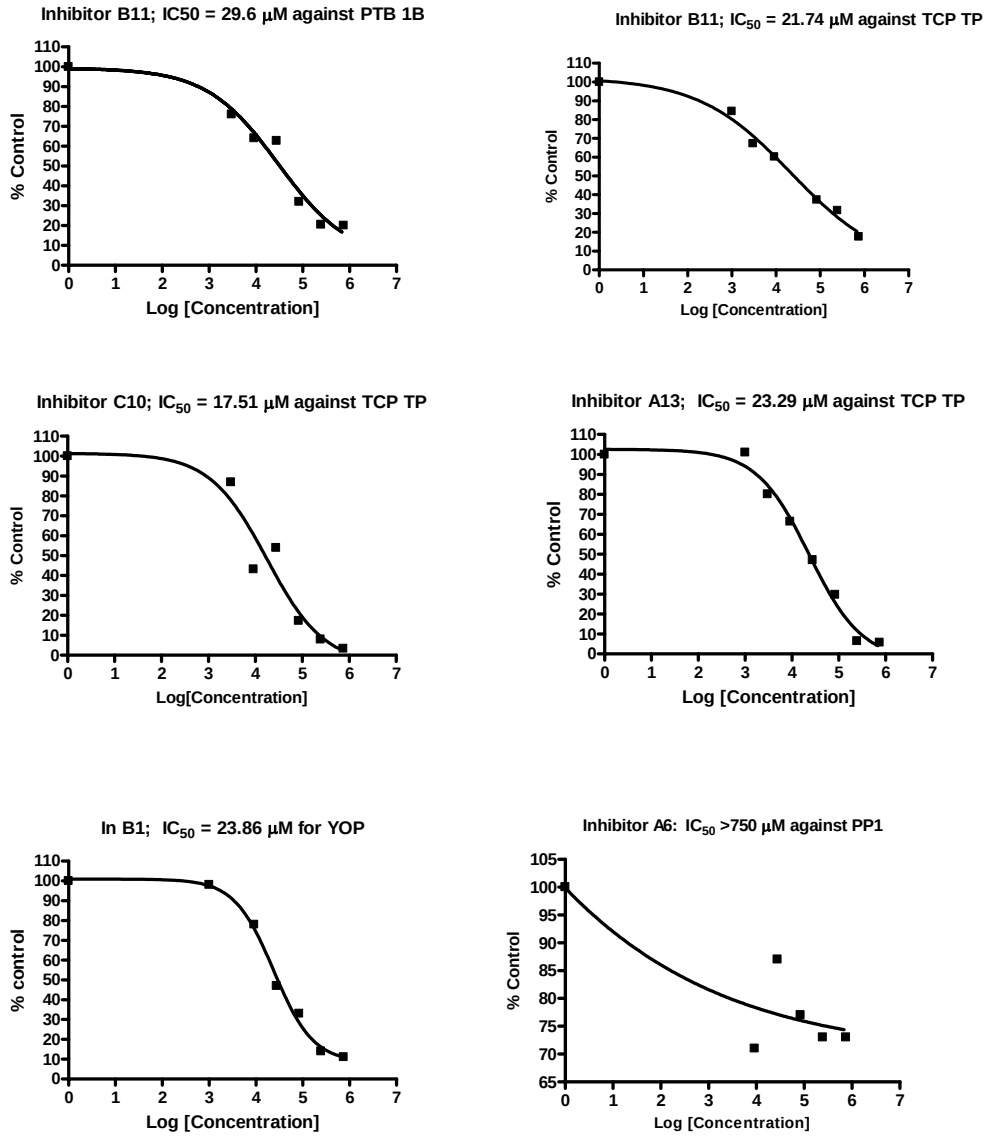


Fig. 7.2: Inhibition profiles of the 66-member library against 6 phosphatases. Note that inhibitor concentrations differ with different enzymes.

IC_{50} measurements of selected inhibitors:

Compounds **2-A13**, **2-C10**, **2-B1**, **2-B11**, **2-B4**, **2-A5** were identified to be the relatively potent ones against PTB1B, TCPTP, YOP, LAR, λPPase, and PP1, respectively. These six compounds were further investigated by the measurement of their IC_{50} values. The IC_{50} values of the six compounds against all six enzymes were shown in previous section. The IC_{50} values were obtained using dose-dependent reactions by varying the concentrations of the inhibitor, under the same enzyme concentration. Briefly, a two-fold dilution series of an inhibitor, from approximately 745 µM to 1 µM (final concentrations) was prepared. The reaction conditions for each enzyme were as listed above. The enzymatic reactions were allowed to incubate at room temperature for

0-60 min before being interrogated for end-point fluorescence at λ_{ex} = 355 nm and λ_{em} = 460 nm. The IC_{50} was calculated by fitting the fluorescence outputs obtained using the Graphpad Prism software v.4.03 (GraphPad, San Diego). Each IC_{50} plots were generated by 6 to 8 data points. Representative results are shown in **Fig. 7.3**.



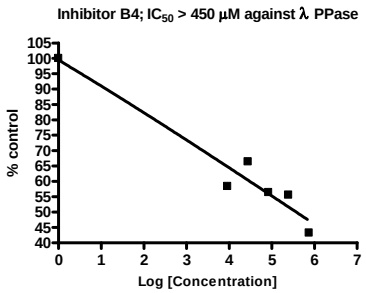


Fig. 7.3 IC_{50} graphs for various 'click' inhibitors against different enzymes.

7.2.3 Chemical synthesis of the 3250-member click library

Synthesis of azides using amine building blocks

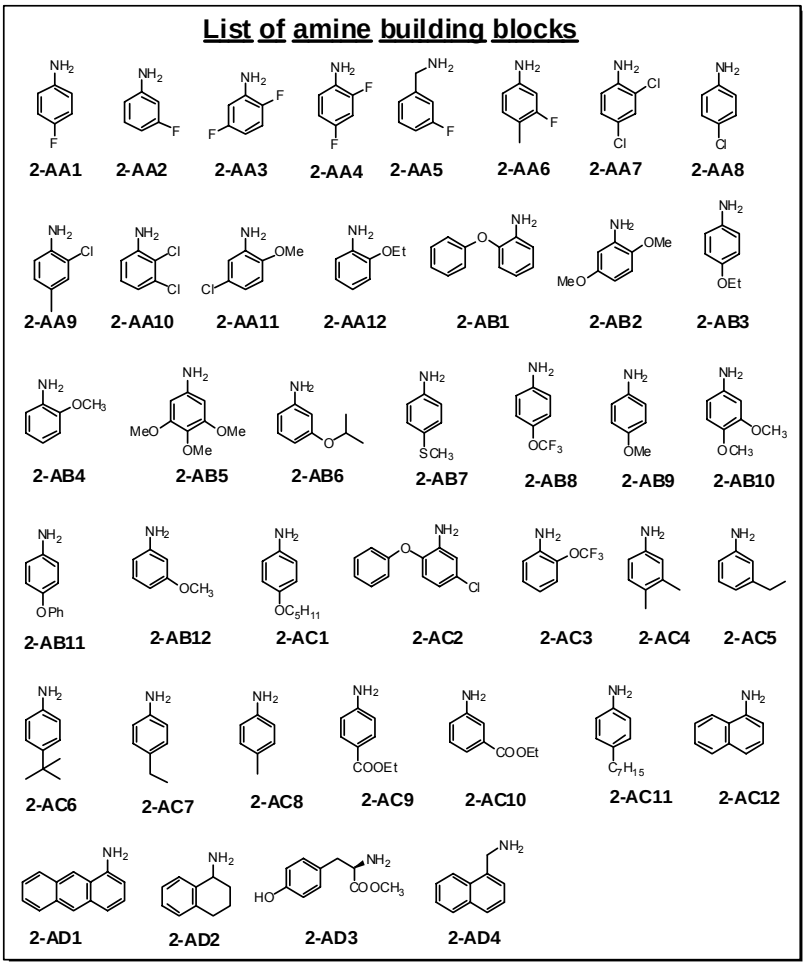


Fig. 7.4 List of amine building blocks

Synthesis of the reductive aminated resin (2-38): PL-FMP resin (200 x 150 mg, 0.9 mmol/g) was taken in 200 MacroKan™ reactors each containing a RF tag. The resin was swelled in 1,2-dichloroethane (250 mL) for about two hours after which the solvent was decanted. The 200 microreactors were then distributed equally into 40 different bottles of capacity 50 mL containing 2% Acetic acid in 1,2-dichloroethane(30 mL). Amines (**AA1-AD4, Fig. 7.4**) (5 eq) were added to the bottles, so that each bottle contains a unique amine. The reaction mixture was then incubated for about 3 hours after which Sodium triacetoxyborohydride (6 eq) was added. After shaken for another 8 hours the solution was decanted and the reactors were combined and washed with DCM (200 mL x 5), MeOH (200 mL x 2) and THF (200 mL x 3) and dried to afford the resin **2-38**.

Synthesis of the N-acylated resin (2-3a to 2-3e): The 200 reductive aminated resins were separated into 5 sets (40 different reductive aminated resins x 5 sets). To the first set, DCM (100 mL), DIEA (10 eq) and acid chloride, **2-i** to **2-l** (5 eq) were added and the reaction mixture was shaken for about 8 hours after which the solution was decanted and the resin was washed with DCM (100 mL x 3), MeOH (100 mL x 2) and THF (100 mL x 3) and dried to afford resin **2-3a** to **2-3e**.

General protocol for cleavage and release of azides (2-A1-2C to 2-D4-6C): Each dried resin was treated with an 1.5 ml solution containing TFA (50%), DCM (45%) and water (5%) and the mixture was shaken for 4 hours and transferred into 3 different 96-well plates and concentrated in vacuo to afford the below azides. Representative yield varied

from 50% to 80%. The azides were redissolved in DMSO (1 mL) to give 50 mM solution (assuming 50 % yield).

All azides synthesized from the above protocols are of high purity (90-95%). Representative compounds were further characterizations, without any purification, by LCMS, ^{1}H and ^{13}C NMR.

2-Azido-N-(3-fluoro-phenyl)-acetamide (2-A2-2C): ^{1}H-NMR (500 MHz, $(CD_3)_2SO$) δ 10.3 (S, 1h), 7.58 (d, *J* = 8 Hz, 1H), 7.39-7.29 (m, 1H), 7.09 (d, *J* = 50 Hz, 1H), 6.92 (m, 1H), 4.06 (s, 2H). ^{13}C-NMR (500 MHz, $(CD_3)_2SO$) δ 167.1, 163.6, 131.0, 140.1, 115.5, 110.6, 106.6, 51.8. ESI-MS(TOF): *m/z* 195.060 $[M + H]^+$.

2-Azido-N-(2-phenoxy-phenyl)-acetamide (2-B1-2C) :^{1}H-NMR (500 MHz, $(CD_3)_2SO$) δ 9.71 (s, 1H), 7.39 (t, ^{3}J = 8 Hz, 3H), 7.16-7.11 (m, 3H), 7.00 (d, ^{3}J = 8Hz, 2H), 6.89-6.87 (m, 1H), 4.06 (s, 2H). ^{13}C-NMR (500 MHz, $(CD_3)_2SO$) δ 156.5, 129.9, 128.9, 125.3, 123.6, 123.5, 118.6, 118.5, 51.1. ESI-MS(TOF): *m/z* 269.098 $[M + H]^+$.

2-Azido-N-(3,4-dimethoxy-phenyl)-acetamide (2-B2-2C) : H-NMR (500 MHz, $(CD_3)_2SO$) δ 9.36 (s, 1H), 7.71 (s, 1H), 6.97 (d, ^{3}J = 8.8 Hz, 1H), 6.66 (m, 1H), 4.14 (s, 2H), 3.79 (s, 3H), 3.69 (s, 3H). ^{13}C-NMR (500 MHz, $(CD_3)_2SO$) δ 166.9, 153.4, 143.9, 127.8, 112.4, 109.0, 108.49, 56.7, 55.8, 51.8. ESI-MS(TOF): *m/z* 237.091 $[M +H]^+$.

2-Azido-N-(4-pentyloxy-phenyl)-acetamide 2-B5-2C) : ^{1}H-NMR (500 MHz, $(CD_3)_2SO$) δ10.06 (s, 1H), 6.96 (s, 2H), 4.00 (s, 2H), 3.74 (s, 6H), 3.62 (s, 3H). ^{13}C-NMR

(500 MHz, $(CD_3)_2SO$) δ 166.0, 152.7, 134.4, 133.6, 97.0, 60.05, 55.6, 51.2. ESI-MS(TOF): *m/z* 267.100 $[M + H]^+$.

2-Azido-N-(4-pentyl-phenyl)-acetamide (2-B12-2C) : ^{1}H-NMR (500 MHz, $(CD_3)_2SO$) δ 10.10 (s, 1H), 7.28 (s, 1H), 7.22 (t, ^3J = 8.2 Hz, 1H), 7.11 (d, ^3J = 8.2 Hz, 1H), 6.66 (d, ^3J = 8 Hz, 1H), 4.02 (s, 2H), 3.73 (s, 3H). ^{13}C-NMR (500 MHz, $(CD_3)_2SO$) δ 166.8, 160.0, 140.0, 130.1, 112.0, 109.6, 105.5, 55.4, 51.7. ESI-MS(TOF): *m/z* 207.080 $[M + H]^+$.

2-Azido-N-(3,4-dimethyl-phenyl)-acetamide (2-C4-2C): H-NMR (500 MHz, $(CD_3)_2SO$) δ 9.96 (s, 1H), 7.36 (s, 1H), 7.29 (d, ^3J = 8.2 Hz, 1H), 7.06 (d, ^3J = 8.2 Hz, 1H), 3.99 (s, 2H), 2.19 (s, 3H), 2.16 (s, 3H). ^{13}C-NMR (500 MHz, $(CD_3)_2SO$) δ 165.8, 136.3, 136.0, 131.4, 129.6, 120.4, 116.7, 51.2, 19.5, 18.7. ESI-MS(TOF): *m/z* 205.101 $[M+ H]^+$.

2-Azido-N-(4-ethyl-phenyl)-acetamide (2-C7-2C): ^{1}H-NMR (500 MHz, $(CD_3)_2SO$) δ 10.0 (s, 1H), 7.47 (d, ^3J = 8.8 Hz, 2H), 7.15 (d, ^3J = 8.2 Hz, 2H), 4.00 (s, 2H), 2.55 (m, 2H), 1.15 (t, ^3J = 7.6 Hz, 3H). ^{13}C-NMR (500 MHz, $(CD_3)_2SO$) δ 166.4, 139.5, 136.6, 128.5, 119.8, 51.7, 28.0, 16.1. ESI-MS(TOF): *m/z* 205.103 $[M +H]^+$.

3-(2-Azido-acetylamino)-benzoic acid ethyl ester (2-C10-2C): ^{1}H-NMR (500 MHz, $(CD_3)_2SO$) δ 10.36 (s, 1H), 8.24 (s, 1H), 7.85 (d, ^3J = 8.2 Hz, 1H), 4.32 (m, ^3J = 7 Hz,

2H), 4.07 (s, 2H), 1.32 (t, ^{3}J = 7 Hz, 3H). ^{13}C-NMR (500 MHz, $(CD_3)_2SO$) δ 166.6, 165.4, 138.7, 130.4, 129.3, 124.2, 123.6, 119.6, 60.8, 51.2, 14.1.

2-Azido-N-(4-heptyl-phenyl)-acetamide (2-C11-2C): ^{1}H-NMR (500 MHz, $(CD_3)_2SO$) δ 10.03 (s, 1H), 7.46 (d, ^{3}J = 8.8 Hz, 2H), 7.12 (d, ^{3}J = 8.8 Hz, 2H), 4.01 (s, 2H), 2.51 (t, 2H), 1.52 (m, 2H), 1.24 (m, 8H), 0.85 (t, ^{3}J = 7 Hz, 3H). ^{13}C-NMR (500 MHz, $(CD_3)_2SO$) δ 165.9, 137.6, 136.0, 128.4, 119.2, 51.2, 34.4, 31.1, 30.9, 28.5, 28.4, 22.0, 13.8. ESI-MS (TOF): *m/z* 275.179 $[M+H]^+$.

N-Anthracen-1-yl-2-azido-acetamide (2-D1-2C): ^{1}H-NMR (500 MHz, $(CD_3)_2SO$) δ 10.40 (s, 1H), 8.48 (d, ^{3}J = 13.8 Hz, 3H), 8.06 (t, ^{3}J = 9.4 Hz, 3H), 7.56 (d, ^{3}J = 9.4 Hz, 1H), 7.47 (m, 2H), 4.14 (s, 2H). ^{13}C-NMR (500 MHz, $(CD_3)_2SO$) δ 167.1, 135.8, 132.2, 131.5, 129.3, 128.5, 128.1, 126.2, 125.6, 125.5, 121.2, 116.7, 114.5, 51.9. ESI-MS (TOF): *m/z* 277.099 $[M + H]^+$.

4-Azido-N-(4-fluoro-phenyl)-butyramide (2-A1-4C): ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.94(s, 1H), 7.57-7.54(m, 2H), 7.09-7.06 (m, 2H), 3.35-3.33(t, *J* = 6.62 Hz, 2H), 2.36-2.33(m, 2H), 1.81-1.78(m, 2H). ^{13}C-NMR (125 MHz, DMSO-d6) δ 170.12, 158.75, 156.84, 135.56, 120.76(d), 115.20(d), 50.24, 33.07, 24.19.

4-Azido-N-(3-fluoro-4-methyl-phenyl)-butyramide (2-A6-4C): ^{1}H-NMR (300 MHz, DMSO-d6) δ 10.03, 7.55-7.52(m, 1H), 7.18-7.17(m, 2H), 3.40-3.37 (t, *J* = 6.625 Hz, 2H), 2.51-2.50(m, 2H), 2.16(s, 3H), 1.86-1.81(q, 2H). ^{13}C-NMR (125 MHz, $CDCl_3$) δ 170.37,

161.12, 159.21, 138.54(d), 131.25(d), 118.18(d), 114.53, 105.80, 5.21, 33.13, 24.52, 13.54.

4-Azido-N-(4-chloro-phenyl)-butyramide (2-A8-4C): ^{1}H-NMR (500 MHz, DMSO-d6) δ 10.02(s, 1H), 7.58-7.56 (d, J = 8.85 Hz, 2H), 7.30-7.28 (d, J = 8.8 Hz, 2H), 3.35-3.33 (t, *J* = 6.95 Hz, 2H), 2.37 (m, 2H), 1.82-1.77(m, 2H). ^{13}C-NMR (125 MHz, DMSO-d6) δ 170.40, 138.09, 128.48, 126.49, 120.52, 50.22, 33.16, 24.12.

4-Azido-N-(4-chloro-2-methoxy-phenyl)-butyramide (2-A11-4C): ^{1}H-NMR (500 MHz, DMSO-d6) δ 9.29(s, 1H), 8.10(s, 1H), 7.10-7.08(m, 1H), 7.05-7.03 (m, 1H), 3.83(s, 3H), 3.39-3.36(t, *J* = 6.92Hz, 2H), 2.54-2.49(m, 2H), 1.86-1.80(m, 2H). ^{13}C-NMR (125 MHz, DMSO-d6) δ 170.94, 148.05, 128.58, 123.74, 123.29, 120.84, 112.85, 112.35, 55.96, 50.21, 32.96, 24.26.

4-Azido-N-(2-methoxy-phenyl)-butyramide (2-B4-4C): ^{1}H-NMR (500 MHz, DMSO-d6) δ 9.06 (s, 1H), 7.88-7.86(m, 1H), 7.03-6.97(m, 2H), 6.86-6.83(t, 1H), 3.8(s, 3H), 3.35-3.32(t, J = 6.92Hz, 2H), 2.46-2.41(m, 2H), 1.80-1.77(m, 2H). ^{13}C-NMR (125 MHz, DMSO-d6) δ 170.41, 149.64, 127.23, 124.26, 122.13, 120.09, 111.06, 55.55, 50.25, 32.94, 24.38.

4-Azido-N-(4-methylsulfanyl-phenyl)-butyramide (2-B7-4C): ^{1}H-NMR (500 MHz, DMSO-d6) δ 9.89(s, 1H), 7.51-7.50(m, 2H), 7.17-1.16(m, 2H), 3.35-3.32 (t, *J* = 6.95 Hz,

3H), 2.38(m,3H), 2.36-2.33 (t, 2H), 1.82-1.77(q, 2H). ^{13}C-NMR (125 MHz, DMSO-d6) δ 170.14, 136.73, 131.42, 127.13, 119.71, 50.26, 33.14, 24.20, 15.57.

4-Azido-N-(4-trifluoromethoxy-phenyl)-butyramide (2-B8-4C): ^{1}H-NMR (500 MHz, DMSO-d6) δ 10.09(s, 1H), 7.66-7.64(d, J = 8.85 Hz, 2H), 7.25-7.23(d, *J* = 8.2 Hz, 2H), 3.36-3.33 (t, *J* = 6.6, 2H), 2.38-2.35(m, 2H), 1.83-1.78(m, 2H). ^{13}C-NMR (125 MHz, DMSO-d6) δ 170.49, 143.34, 138.37, 121.47, 120.30, 119.09, 50.24, 33.16, 24.15.

4-Azido-N-(4-phenoxy-phenyl)-butyramide (2-B11-4C): ^{1}H-NMR (500 MHz, DMSO-d6) δ 9.92(s, 1H), 7.57-7.56(m, 2H), 7.32-7.29 (m, 2H), 7.05-7.02(m, 1H), 6.94-6.90(m, 4H), 3.36-3.33(t, *J* = 6.92 Hz, 2H), 2.37-2.34(t, *J* = 7.25 Hz, 2H), 1.83-1.78 (m, 2H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 170.04, 157.38, 151.47, 135.13, 129.85, 122.83, 120.70, 119.37, 117.72, 50.26, 33.10, 24.26.

4-Azido-N-(4-methoxy-phenyl)-butyramide (2-B12-4C): ^{1}H-NMR (500 MHz, DMSO-d6) δ 9.87(s, 1H), 3.67 (s, 3H), 3.36-3.329(m, 2H), 2.36-2.33 (t, *J* = 7.25 Hz, 2H), 1.82-1.77(m, 2H). ^{13}C-NMR (125 MHz, DMSO-d6) δ 170.31, 159.44, 140.34, 129.36, 111.32, 108.42, 104.86, 54.86, 50.25, 33.23, 24.18.

4-Azido-N-(4-tert-butyl-phenyl)-butyramide (2-C6-4C): ^{1}H-NMR (500 MHz, DMSO-d6) δ 9.80(s, 1H), 7.45-7.44(d, *J* = 8.8 Hz, 2H), 7.25-7.24(d, J = 8.2 Hz, 2H), 3.35-3.32(t, *J* = 6.62 Hz, 2H), 2.35-2.32 (t, *J* = 7.57 Hz, 2H), 1.82-1.76(m, 2H), 1.2 (s, 9H). ^{13}C-NMR

(75 MHz, DMSO-d6) δ 170.00, 145.26, 136.58, 125.14, 118.84, 50.27, 33.89, 33.10, 31.13, 24.28

4-Azido-N-(4-ethyl-phenyl)-butyramide (2-C7-4C): ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.36-7.33(d, *J* = 8.7 Hz, 2H), 6.85-6.82(d, *J* = 8.7 Hz, 2H), 4.03-3.96(m, 2H), 3.42-3.38(t, *J* = 6.27 Hz, 2H), 2.47-2.43 (t, *J* = 7.14 Hz, 2H), 2.01-1.97 (t, *J* = 6.79Hz, 2H), 1.41-1.37(t, J = 6.79 Hz, 3H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 170.83, 142, 130, 122.30, 114.76, 63.68, 50.64, 33.91

5-Azido-pentanoic acid (2-ethoxy-phenyl)-amide (2-A12-5C): ^{1}H-NMR (500 MHz, DMSO-d6) δ 8.87(s, 1H), 7.85(m, 1H), 7.00-6.97(m, 2H), 6.86-6.83(m, 1H), 4.06-3.99(m, 2H), 3.34-3.32(m, 2H), 2.52-2.37(m, 2H), 1.58(m, 4H), 1.32-1.21(m, 3H). ^{13}C-NMR (125 MHz, $CDCl_3$) δ 170.87, 148.98, 127.45, 124.33, 122.37, 120.07, 112.17, 63.82, 50.38, 44.75, 35.40, 27.74, 22.39, 14.52. ESI-MS (TOF): *m/z* 263.148 $[M + H]^+$.

5-Azido-pentanoic acid (4-ethoxy-phenyl)-amide (2-B3-5C): ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.68 (s, 1H), 7.44-7.41 (d, *J* = 15 Hz, 2H), 6.81-6.78(d, *J* = 15.05 Hz, 2H), 3.95-3.88(m, 2H), 3.33-3.28(d, *J* = 10.67 Hz, 2H), 2.27-2.23 (t, *J* = 11.5 Hz, 2H), 1.58-1.52(m, 4H), 1.27-1.20(m, 3H). ^{13}C-NMR (125 MHz, DMSO-d6) δ 170.27, 154.24, 132.33, 120.55, 114.26, 63.01, 50.38, 35.55, 27.84, 22.43, 14.64. ESI-MS(TOF): *m/z* 263.148 $[M + H]^+$.

5-Azido-pentanoic acid (2-methoxy-phenyl)-amide (2-B4-5C): ^{1}H-NMR (300 MHz, DMSO-d6) δ 8.99(s, 1H), 7.89-7.87 (d, *J* = 12.85, 1H), 7.04-6.96 (m, 2H), 6.87-6.81(m, 1H), 3.77(s, 3H), 3.33-3.29 (t, *J* = 10.55 Hz, 2H), 2.39-2.35(t, *J* = 11.1 Hz, 2H,), 1.61-1.51(m, 4H). ^{13}C-NMR (125 MHz, DMSO-d6) δ 170.99, 149.60, 127.30, 124.20, 122.06, 120.10, 111.04, 55.55, 50.35, 35.35, 27.77, 22.37. ESI-MS(TOF): *m/z* 249.133 $[M + H]^+$.

5-Azido-pentanoic acid (3-isopropoxy-phenyl)-amide (2-B6-5C): ^{1}H-NMR (500 MHz, DMSO-d6) δ 9.78 (s, 1H), 7.24 (s, 1H), 7.13-7.08(m, 2H), 6.54-6.51(m, 1H), 4.51-4.43(m, 1H), 3.3 (t, 2H), 2.30-2.26(t, *J* = 11.65Hz, 2H), 1.63-1.50(m, 4H), 1.21 (s, 3H), 1.23 (s, 3H). ^{13}C-NMR (125 MHz, $CDCl_3$) δ 170.85, 157.64, 140.42, 129.34, 111.11, 110.12, 106.52, 69.06, 50.38, 35.74, 27.81, 22.23, 21.79. ESI-MS(TOF): *m/z* 277.163 $[M + H]^+$.

5-Azido-pentanoic acid (3,4-dimethoxy-phenyl)-amide (2-B10-5C): ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.7 (s, 1H), 7.06-7.03(d, J = 8.55Hz, 1H), 6.83-6.80(d, *J* = 8.52 Hz, 1H), 3.7(s, 6H), 3.33-3.29(t, *J* = 6.24 Hz, 2H), 2.28-2.24(t, J = 6.73 Hz, 2H), 1.60-1.50(m, 4H). ^{13}C-NMR (75 MHz, DMSO-d6) δ 170.39, 148.50, 144.63, 132.96, 112.08, 110.94, 104.33, 55.71, 55.30, 50.40, 38.63, 35.65, 27.86, 22.32. ESI-MS(TOF): *m/z* 279.143 $[M + H]^+$.

5-Azido-pentanoic acid (5-chloro-2-phenoxy-phenyl)-amide (2-C2-5C): ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.56(s, 1H), 8.10(s, 1H), 7.37-7.32(m, 2H), 7.10-7.06(m, 2H), 6.97-6.94 (d, *J* = 13.95 Hz, 2H), 6.87-6.84(d, *J* = 14.5 Hz, 1H), 3.24 (t, 2H), 2.36-2.31 (t,

J = 11.5 Hz, 2H), 1.52-1.41(m, 4H). ^{13}C-NMR (75 MHz, DMSO-d6) δ 171.67, 156.34, 145.81, 131.15, 129.89, 127.13, 124.11, 123.59, 122.71, 120.09, 118.30. ESI-MS(TOF): *m/z* 345.110 $[M + H]^+$.

5-Azido-pentanoic acid (3-ethyl-phenyl)-amide (2-C5-5C): ^{1}H-NMR (500 MHz, DMSO-d6) δ 9.77 (s, 1H), 7.40-7.31 (m, 2H), 7.16-7.07(m, 1H), 6.83-6.77 (m, 1H), 3.33-3.27(m, 2H), 2.52(m, 2H), 2.30-2.22(m, 2H), 1.55(m, 4H), 1.14-1.05(m, 3H). ^{13}C-NMR (125 MHz, $CDCl_3$) δ 170.77, 144.12, 139.24, 128.47, 122.44, 118.39, 116.49, 50.39, 35.70, 28.23, 27.83, 22.29, 15.46. ESI-MS(TOF): *m/z* 247.153 $[M + H]^+$

4-(5-Azido-pentanoylamino)-benzoic acid ethyl ester (2-C9-5C): ^{1}H-NMR (500 MHz, DMSO-d6) δ 10.19(s, 1H), 7.87-7.84(d, *J* = 12.35 Hz, 2H), 7.70-7.67(d, *J* = 13.6 Hz, 2H), 4.26-4.22(m, 2H), 3.33-3.29(t, *J* = 10 Hz, 2H), 2.36-2.32(t, *J* = 11.1 Hz, 2H), 1.61-1.51(m, 4H), 1.28-1.22(m, 3H). ^{13}C-NMR (125 MHz, $CDCl_3$) δ 171.48, 165.29, 143.56, 130.14, 123.95, 118.31, 60.32, 50.38, 35.78, 27.79, 22.10, 14.15. ESI-MS(TOF): *m/z* 291.144 $[M + H]^+$.

6-Azido-hexanoic acid (4-fluoro-phenyl)-amide (2-A1-6C): ^{1}H-300 MHz, DMSO d6) δ 9.87 (s, 1H, NH), 7.55 (m, 2H), 7.07 (m, 2H), 3.28 (t, 2H, *J* = 6.74 Hz), 2.25 (t, 2H, *J* = 7.23 Hz), 1.51 (m, 4H), 1.31 (m, 2H); ^{13}C-NMR (75.5 MHz, DMSO d6) δ 171.31, 159.74, 136.06, 121.13, 115.37, 50.89, 36.49, 28.40, 26.18, 24.97. ESI-MS (TOF): m/*z* = $[M+H]^+$ 251.124

6-Azido-hexanoic acid (4-chloro-phenyl)-amide (2-A8-6C): ^{1}H-NMR (300 MHz, DMSO d6) δ 9.95 (s, 1H, NH), 7.56 (d, 2H, *J* = 8.85 Hz), 7.30 (d, 2H, *J* = 8.88 Hz), 3.28 (t, 2H, *J* = 6.82 Hz), 2.27 (t, 2H, *J* = 7.30 Hz), 1.51 (m, 4H), 1.33 (m, 2H); ^{13}C-NMR (75.5 MHz, DMSO d6) δ 171.59, 138.62, 128.89, 126.81, 120.89, 50.88, 36.57, 28.39, 26.15, 24.90. ESI-MS(TOF): m/z = $[M+H]^{+}$ 267.091

6-Azido-hexanoic acid (4-chloro-2-methoxy-phenyl)-amide (2-A11-6C): ^{1}H- NMR (300 MHz, DMSO d6) δ 9.15 (s, 1H, NH), 8.06 (s, 1H), 7.10-6.95 (m, 2H), 3.79 (s, 3H, OMe), 3.28 (t, 2H, *J* = 6.80 Hz), 2.37 (t, 2H, J = 7.32 Hz), 1.52 (m, 4H), 1.31 (m, 2H); ^{13}C-NMR (75.5 MHz, DMSO d6) δ 172.10, 148.40, 129.09, 124.14, 123.54, 121.11, 112.76, 56.38, 50.91, 36.19, 28.39, 26.11, 24.99. ESI-MS(TOF): m/z = $[M+H]^{+}$ 297.104

6-Azido-hexanoic acid (2-ethoxy-phenyl)-amide (2-A12-6C): ^{1}H- NMR (300 MHz, DMSO d6) δ 8.81 (s, 1H, NH), 7.85 (d, 1H, *J* = 7.74 Hz), 7.03-6.95 (m, 2H), 6.86-6.79 (m, 1H), 4.03 (q, 2H, *J* = 6.9 Hz), 3.29 (t, 2H, *J* = 6.74 Hz), 2.34 (t, 2H, J = 7.32 Hz), 1.52 (m, 4H), 1.32 (m, 2H); ^{13}C-NMR (75.5 MHz, DMSO d6) δ 171.46, 149.31, 127.93, 124.64, 122.67, 120.47, 112.59, 64.22, 50.93, 36.31, 28.42, 26.12, 25.12, 14.93. ESI-MS (TOF): m/z = $[M+H]^{+}$ 277.157

6-Azido-hexanoic acid (4-ethoxy-phenyl)-amide (2-B3-6C): ^{1}H- NMR (300 MHz, DMSO d6) δ 9.64 (s, 1H, NH), 7.42 (d, 2H, *J* = 9.03 Hz), 6.79 (d, 2H, *J* = 8.88 Hz),), 3.92 (q, 2H, *J* = 7.08 Hz), 3.28 (t, 2H, *J* = 6.82 Hz), 2.22 (t, 2H, *J* = 7.39 Hz), 1.52 (m, 4H), 1.26 (m, 2H); ^{13}C-NMR (75.5 MHz, DMSO d6) δ 170.86, 154.60, 132.79, 120.92,

114.66, 63.41, 50.90, 36.45, 28.41, 26.20, 25.07, 15.04. ESI-MS(TOF): m/z = $[M+H]^+$ = 7.156

6-Azido-hexanoic acid (2-methoxy-phenyl)-amide (2-B4-6C): ^{1}H- NMR (300 MHz, DMSO d6) δ 8.97 (s, 1H, NH), 7.88 (d, 1H, *J* = 7.71 Hz), 6.99 (m, 2H), 6.84 (m, 1H), 3.78 (s, 3H, OMe), 3.28 (t, 2H, *J*= 6.82 Hz), 2.34 (t, 2H, J = 7.23 Hz), 1.55 (m, 4H), 1.34 (m, 2H); ^{13}C-NMR (75.5 MHz, DMSO d6) δ 171.59, 149.96, 127.77, 124.53, 122.40, 120.51, 111.44, 55.97, 50.94, 36.24, 28.41, 26.16, 25.12. ESI-MS(TOF): m/z = $[M+H]^+$ = 3.142

6-Azido-hexanoic acid (3,4,5-trimethoxy-phenyl)-amide (2-B5-6C): ^{1}H- NMR (300 MHz, DMSO d6) δ 9.74 (s, 1H, NH), 6.94 (s, 2H), 3.68 (s, 6H, OMe), 3.58 (s, 3H, OMe), 3.29 (t, 2H, *J* = 6.73 Hz), 2.23 (t, 2H, *J* = 7.23 Hz),1.53 (m, 4H), 1.31 (m, 2H); ^{13}C-NMR (75.5 MHz, DMSO d6) δ 171.29, 153.04, 135.86, 133.56, 97.13, 60.45, 56.02, 50.90, 36.65, 28.43, 26.19, 24.98. ESI-MS(TOF): m/z = $[M+H]^+$ 323.160

6-Azido-hexanoic acid (3,4-dimethoxy-phenyl)-amide (2-B10-6C): ^{1}H- NMR (300 MHz, DMSO d6) δ 9.67 (s, 1H, NH), 7.26 (s, 1H), 7.03 (d, 1H, *J* = 8.55 Hz), 6.82 (d, 1H, *J* = 8.70 Hz), 3.68 (s, 6H, OMe), 3.29 (t, 2H, *J* = 6.73 Hz), 2.23 (t, 2H, *J* = 7.30 Hz),1.53 (m, 4H), 1.31 (m, 2H); ^{13}C-NMR (75.5 MHz, DMSO d6) δ 170.96, 148.88, 144.97, 133.40, 112.49, 111.28, 104.69, 56.12, 55.69, 50.90, 36.54, 28.42, 26.20, 25.03. ESI-MS: m/z = $[M+H]^+$ 293.153

6-Azido-hexanoic acid (4-phenoxy-phenyl)-amide (2-B11-6C): ^{1}H- NMR (300 MHz, DMSO d6) δ 9.84 (s, 1H, NH), 7.55 (d, 2H, *J* = 8.88 Hz), 7.31 (t, 2H, *J* = 7.56 Hz), 7.05 (t, 1H, *J* = 7.39 Hz), 6.92 (m, 4H), 3.28 (t, 2H, *J* = 6.80 Hz), 2.26 (t, 2H, *J* = 7.32 Hz),1.54 (m, 4H), 1.34 (m, 2H); ^{13}C-NMR (75.5 MHz, DMSO d6) δ 171.22, 157.82, 151.79, 135.66, 130.28, 123.23, 121.04, 119.81, 118.11, 50.90, 36.51, 28.41, 26.18, 25.03. ESI-MS(TOF): m/z = $[M+H]^{+}$ 325.159

Synthesis of azides using acid building blocks

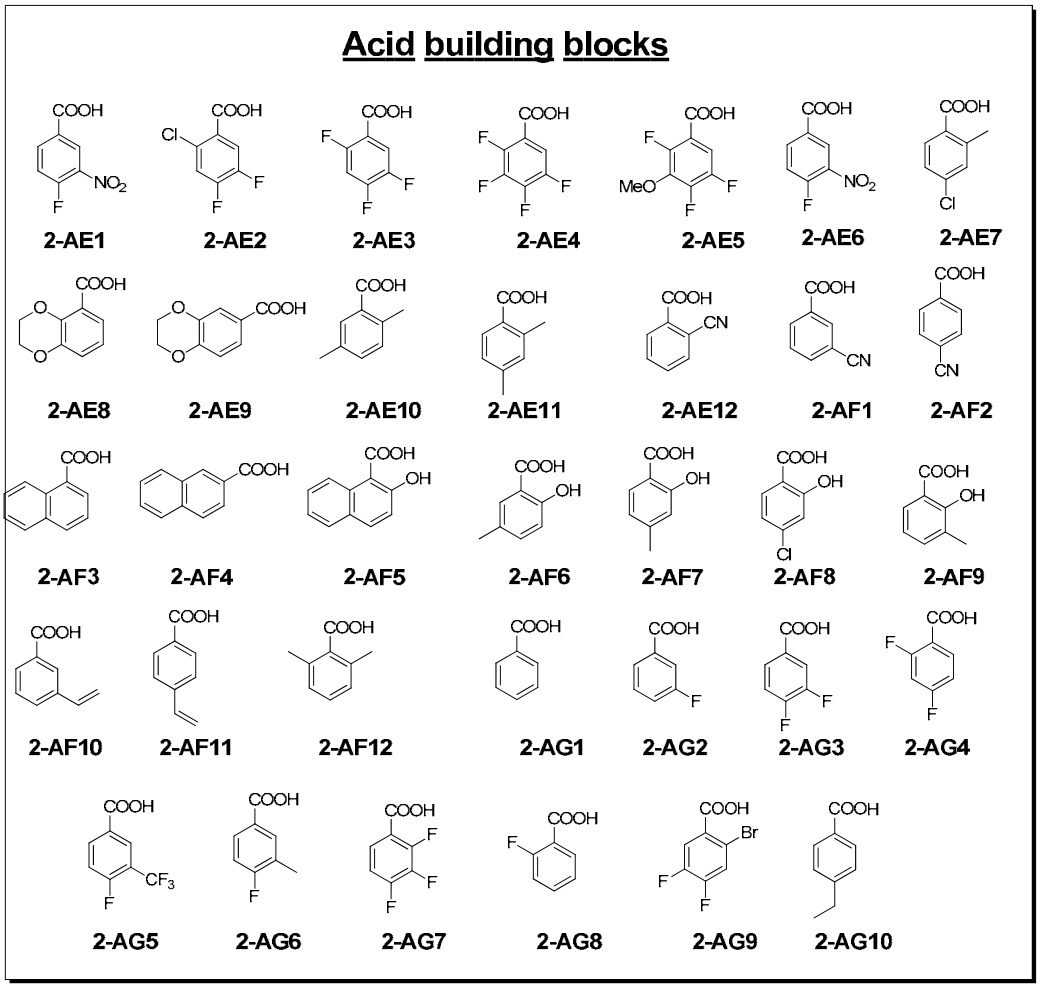


Fig. 7.5 List of acid building blocks.

Synthesis of the reductive aminated resin (2-4a to 2-4e):

Pre-swelled PL-FMP resin (34 x 150 mg, 0.9 mmol/g) was taken in to 5 sets (x 34) MacroKan™ reactors each containing a RF tag. The reactors were taken into a 250 ml bottle containing 2% Acetic acid in 1,2-dichloroethane (150 mL), DIEA (8 eq) and the amine **6a-e** (4 eq). After incubating for about 2 hours, Sodium triacetoxyborohydride (5 eq) was added. After shaken for another 4 hours, the solution was decanted and the microreactors were washed with DCM (200 mL x 5), MeOH (200 mL x 2) and THF (200 mL x 3) and dried to afford the N-acylated resins **2-4a** to **2-4e** (5 sets, each contains 34 identical reactors).

Synthesis of N-acylated resin (2-39):

To the first set of the reductive aminated resin **2-4a** to **2-4e** (34 x 150 mg) was taken into 34 different bottles (10 mL). DMF solution (5 mL) containing PyBOP (4 eq), HOAt (4 eq) and DIEA (6 eq) was added followed by the addition of acid building blocks **2-AE1-2AG10** (4 eq), **(Fig. 7.5)** so that each bottle had a unique acid. After shaken for 12 hours the microreactors were decanted and combined and washed with DCM (200 mL x 5), MeOH (200 mL x 2) and THF (200 mL x 3) and dried to afford the resin **2-39**.

Cleavage and release of azides (2-E1-2C to 2-G10-2C):

The above General protocol for cleavage and release of azides, to afford the **2-E1-2C** to **2-G10-2C**. All azides synthesized from the above protocols are of high purity (90-95%) Representative compounds were further characterizations, without any purification, by LCMS, ^{1}H and ^{13}C NMR.

N-(2-Azido-ethyl)-benzamide (2-G1-3C): ^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.79-7.77 (d, *J* = 12.20 Hz, 2H), 7.54-7.41 (m, 4H), 3.64-3.56 (m, 4H). ^{13}C-NMR (125 MHz, $CDCl_3$) 132.65, 132.29, 129.21, 127.53, 51.50, 40.00. ESI-MS(TOF): *m/z* 191.092 $[M + H]^+$.

N-(2-Azido-ethyl)-3-fluoro-benzamide (2-G2-3C): ^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.57-7.51(m, 2H), 7.43-7.36(q, 4H), 7.22-7.17(m, 1H), 3.62-3.55(m, 4H). ^{13}C-NMR (125 MHz, $CDCl_3$) 167.19, 136.92, 136.92, 130.84(d), 123.12(d), 119.32(d), 115.13(d), 51.22, 40.08. ESI-MS(TOF): *m/z* 209.081 $[M + H]^+$.

N-(2-Azido-ethyl)-3,4-difluoro-benzamide (2-G3-3C): ^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.73-7.66(m, 1H), 7.58-7.55(m, 1H), 7.28-7.17(m, 1H), 3.62-3.55(m, 4H). ^{13}C-NMR (125 MHz, $CDCl_3$) 165.77, 154.19, 151.98, 131.17, 123.17(m), 117.61(m), 51.10, 40.85. ESI-MS(TOF): *m/z* 227.073 $[M + H]^+$.

N-(2-Azido-ethyl)-2,4-difluoro-benzamide (2-G4-3C): ^{1}H-NMR (500 MHz, $CDCl_3$) δ 8.12-8.10(d, *J* = 9.85 Hz, 1H), 3.61-3.56(m, 4H). ^{13}C-NMR (125 MHz, $CDCl_3$) 166.09, 164.0, 161, 133.77(d), 127.26, 117.94(d), 51.13, 40.20. ESI-MS(TOF): *m/z* 227.072 $[M + H]^+$.

N-(2-Azido-ethyl)-4-fluoro-3-trifluoromethyl-benzamide (2-G5-3C): ^{1}H-NMR (500 MHz, $CDCl_3$) δ 8.14-8.10 (m, 1H), 7.02-6.94 (m, 1H), 6.90-6.86(m, 1H), 3.67-3.63(m, 2H), 3.58-3.55(t, *J* = 5.67 Hz, 2H). ^{13}C-NMR (125 MHz, $CDCl_3$) 166.49, 163.24(d),

134.44(m), 117.70(d), 113.10(m), 105.12(m), 51.31, 39.97. ESI-MS (TOF): *m/z* 227.068 $[M + H]^+$.

N-(2-Azido-ethyl)-4-fluoro-3-methyl-benzamide (2-G6-3C): ^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.66(m, 1H), 7.59-7.57(m, 1H), 7.04-7.00(t, *J* = 10 Hz , 1H), 3.62-3.58(m, 2H), 3.55-3.53(t, *J* = 5 , 2H), 2.40 (s, 3H). ^{13}C-NMR (125 MHz, $CDCl_3$) 167.60, 164.94, 162.94, 131.35, 51.42, 40.03, 15.04.

N-(2-Azido-ethyl)-2,3,4-trifluoro-benzamide (2-G7-3C): ^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.87-7.82(m, 1H), 7.12-7.07 (m, 1H), 3.67-3.64(m, 2H), 3.59-3.56(t, *J* = 5.67 Hz, 2H). ESI-MS(TOF): *m/z* 245.060 $[M + H]^+$.

N-(2-Azido-ethyl)-2-fluoro-benzamide (2-G8-3C): ^{1}H-NMR (500 MHz, $CDCl_3$) 8.09-8.06 (m, 1H), 7.50-7.46(m, 1H), 7.27(m, 1H), 7.15-7.11 (m, 1H), 3.68-3.64(m, 2H), 3.58-3.56 (t, *J* = 5.67, 2H). ^{13}C-NMR (125 MHz, $CDCl_3$) δ 164.22, 160.26, 134.16(d), 132.56(d), 125.43(d), 121.22, 116.74(d), 51.33, 39.94. ESI-MS(TOF): *m/z* 209.081 $[M + H]^+$.

N-(2-Azido-ethyl)-4-ethyl-benzamide (2-G10-3C): ^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.72-7.69(d, *J* = 13.9 Hz, 2H), 7.25-7.22 (d, *J* = 13.9 Hz, 2H), 3.63-3.50(m, 4H), 2.72-2.64(q, 2H), 1.26-1.21(t, *J* = 12.7 Hz, 3H). ^{13}C-NMR (125 MHz, $CDCl_3$) δ 168.43, 148.94, 131.94, 128.61, 127.62, 51.42, 39.93, 29.28, 15.79. ESI-MS (TOF): *m/z* 304.983 $[M + H]^+$.

Synthesis of azides using acid chloride building block

Synthesis of N-acylated resin from acid chloride (2-6a to 2-6e):

The reductive aminated resins **2-4a**; **2-4c**; **2-4d**; **2-4e** (4 x 150 mg) were taken in a bottle (25 ml) containing DCM (15 ml). 2-chlorobenzoyl chloride, **H** (4 eq) and DIEA (5 eq) were added. After shaken for 6 hours, the solution was decanted and the microreactors were washed with DCM (50 mL x 5), MeOH (50 mL x 2) and THF (50 mL x 3) and dried to afford the N-acylated resin **2-6a** to **2-6e**

Cleavage and release of the azides (2-H1-2C to 2-H1-6C):

Each dried resin was treated with an 1.5 ml solution containing of TFA (10 %) and DCM (90 %) the mixture was shaken for 1 hour and transferred into 96-well plate, dried *in vacuo*, then redissolved in DMSO (1 mL) to give 50 mM solutions (assuming 50 % yield). All azides synthesized from the above protocols are of high purity (90-95%). Representative compounds were further characterizations, without any purification, by LCMS, ^{1}H and ^{13}C NMR.

Synthesis of azides using sulfonyl chlorides

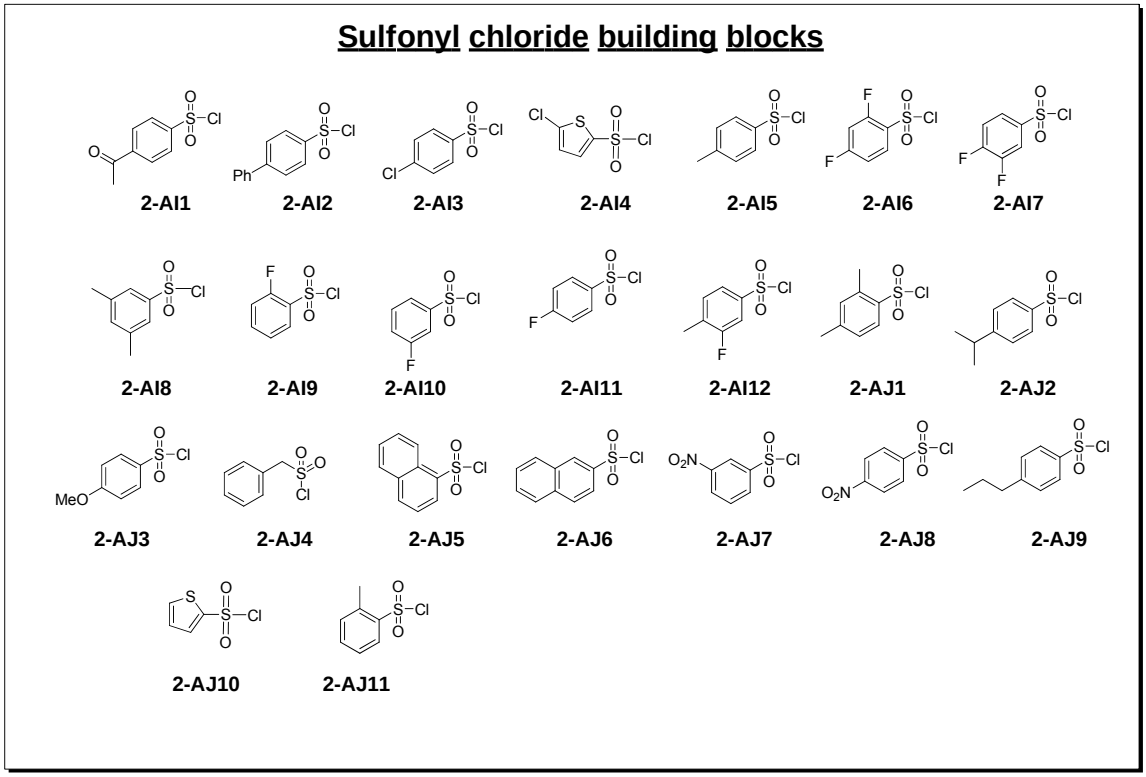


Fig. 7.6 List of Sulfonyl chloride building blocks

Synthesis of the N-sulfonamide resin (2-7a to 2-7-e):

The reductive aminated resin **2-4a** to **2-4e** (23 x 150 mg) was taken in to 23 different bottles (10 ml) each containing a unique Sulfonyl chloride (**2-AI1** to **a-AJ11**) (**Fig. 7.6**) and DIEA (5 eq) in DCM (5 ml). After shaken for 6 hours, the solution was decanted and the microreactors were washed with DCM (200 mL x 5), MeOH (200 mL x 2) and THF (200 mL x 3) and dried to afford the N-sulfonylated resin **2-7a** to **2-7-e**.

Cleavage and release of the sulfonamide-based azides (2-I1-2C to 2-J11-6C):

Each dried resin was treated with an 1.5 ml solution containing of TFA (10 %) and DCM (90 %) the mixture was shaken for 1 hour and transferred into 2 different 96-well plates, dried *in vacuo*, then redissolved in DMSO (1 mL) to give 50 mM solutions (assuming 50 % yield). All azides synthesized from the above protocols are of high purity (90-95%). Representative compounds were further characterizations, without any purification, by LCMS, ^{1}H and ^{13}C NMR.

4-Acetyl-*N*-(2-azidoethyl)benzenesulfonamide (2-I1-2S): ^{1}H NMR (300 MHz, DMSO-d_6) δ 2.63 (s, 3H), 2.94-3.00 (m, 2H), 3.33 (t, J = 5.8 Hz, 2H), 7.93 (d, J = 8.4 Hz, 2H), 8.13-8.16 (m, 3H). ^{13}C NMR (75 MHz, DMSO-d6) δ 27.0, 42.1, 50.1, 126.8, 129.0, 139.5, 144.0, 197.3.

Synthesis of azides using chloroformate building block

Synthesis of carbamate resin (2-8a, 2-8d, 2-8d)

The reductive aminated resins **2-4a**, **2-4d, 2-4e** (3 x 150 mg) were taken in a bottle (25 ml) containing DCM (15 ml). Benzylchloro formate and DIEA (5 eq) were added. After shaken for 6 hours, the solution was decanted and the microreactors were washed with DCM (50 mL x 5), MeOH (50 mL x 2) and THF (50 mL x 3) and dried to afford the carbamate resin **2-8a, 2-8d, 2-8d**

Cleavage and release of the azides (2-L1-2C to 2-L1-6C):

Each dried resin was treated with an 1.5 ml solution containing of TFA (10 %) and DCM (90 %) the mixture was shaken for 1 hour and transferred into 96-well plate, dried *in vacuo*, then redissolved in DMSO (1 mL) to give 50 mM solutions (assuming 50 % yield). All azides synthesized from the above protocols are of high purity (90-95%; see Appendix for detailed structures, ID and characterizations). Representative compounds were further characterizations, without any purification, by LCMS, ^{1}H and ^{13}C NMR.

General procedure for the synthesis of various Propargyl phenyl ethers (2-42, 2-43, 2-49, 2-50, 2-51): Potassium carbonate (99 mmol) and benzo-18-crown-6 (3.3 mmol) were added to a solution of propargyl p-toluenesulfonate (66 mmol) and hydroxyl acetophenone (66 mmol) in acetonitrile (100 ml) followed by refluxing for 6 h, after which the organic phase was removed under reduced pressure and taken into dichloromethane layer (80 ml) and extracted with $NaHCO_3$ (2 x 40 ml), water (2 x 40 ml) and brine (1 x 40 ml). The organic phase was then dried with anhydrous Na_2SO_4, and the solvent was removed *in vacuo* to afford the crude product, which, upon further purification by flash column chromatography, afforded the pure Propargyl phenyl ether. Compounds **2-42, 2-43, 2-49, 2-50, 2-51** were obtained from compounds **2-40, 2-41, 2-46, 2-47, 2-48** respectively, in 67–90% yield.

1-(3,5-dibromo-2-(prop-2-ynyloxy)phenyl)ethanone (2-42): Pale yellow solid. Yield = 67 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.84 (d, J = 2.46 Hz, 1H), 7.71 (d, J = 2.46 Hz, 1H), 4.76 (d, J = 2.43 Hz, 2H), 2.67 (s, 3H), 2.53 (t, J = 2.43 Hz, 1H); ESI-MS: *m/z* $[M-H]^-$ = 331.0.

1-(4-fluoro-2-(prop-2-ynyloxy)phenyl)ethanone (2-43): Yellow solid. Yield = 88%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.85 – 7.79 (m, 1H), 6.82 – 6.72 (m, 2H), 4.80 (d, J = 2.46 Hz, 2H), 2.61 – 2.59 (m, 4H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 197.6, 167.6, 164.2, 158.5, 132.7 (J_{CF} = 10.9 Hz), 125.0 (J_{CF} = 3.3 Hz), 108.6 (J_{CF} = 21.3 Hz), 101.0 (J_{CF} = 26.2 Hz), 76.7, 56.5, 31.8; ESI-MS: *m/z* $[M-H]^-$ = 191.2.

1-(3-methoxy-4-(prop-2-ynyloxy)phenyl)ethanone (2-49): White solid. Yield = 90%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.57 – 7.53 (m, 2H), 7.04 (d, J = 8.22 Hz, 1H), 4.82 (d, J = 2.46 Hz, 2H), 3.92 (s, 3H), 2.59 – 2.57 (m, 4H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 196.7, 150.8, 149.4, 131.4, 122.7, 112.3, 110.5, 76.5, 76.4, 56.5, 55.9, 26.1 ESI-MS: *m/z* $[M+H]^+$ = 205.2.

1-(2-(prop-2-ynyloxy)phenyl)ethanone (2-50): Pale yellow solid. Yield = 63 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.81 – 7.73 (m, 1H), 7.50 – 7.44 (m, 1H), 7.08 – 7.02 (m, 2H), 4.81 (d, J = 2.46 Hz, 2H), 2.64 (s, 3H), 2.54 (t, J = 2.46 Hz, 1H); ESI-MS: *m/z* $[M+Na]^+$ = 197.0.

General procedure for the synthesis of various isoxazole-3-carboxylic acid methyl esters (2-44, 2-45, 2-52, 2-53, 2-54):

To a mixture of alkyne-derivatized acetophenone (53 mmol) and dimethyl oxalate (53 mmol) was added freshly prepared NaOMe (0.5 M in MeOH, 53 mmol). The reaction mixture was refluxed for 12 h before cooled to room temperature. To the same reaction

MeOH (80 ml), $NH_2OH.HCl$ (53 mmol) and a catalytic amount of *p*-$TsOH.H_2O$ were added and the resulting mixture was refluxed continuously for 3 days. Upon cooling to room temperature, the precipitated compound was collected, washed with water and ice-cold methanol to afford the pure isoxazole-3-carboxylic acid methyl ester**.** Compounds **2-44, 2-45, 2-52, 2-53, 2-54** were obtained from compounds **2-42, 2-43, 2-49, 2-50, 2-51** respectively, in 34–44% yield.

Methyl 5-(3,5-dibromo-2-(prop-2-ynyloxy)phenyl)isoxazole-3-carboxylate (2-42): Pale brown powder. Yield = 35%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 8.12 (d, *J* = 2.13 Hz, 1H), 8.07 (d, *J* = 2.31 Hz, 1H), 7.37 (s, 1H), 4.81 (d, *J* = 2.13 Hz, 2H), 3.94 (s, 3H), 3.61 (t, *J* = 2.30 Hz, 1H); ^{13}C-NMR (75 MHz, DMSO-d6) δ 165.7, 159.6, 156.6, 150.6, 137.6, 129.8, 124.1, 119.6, 118.2, 105.1, 80.1, 77.7, 60.6, 53.0; ESI-MS: *m/z* = N.A. (didn't show up in both +ve and –ve modes, probably due to difficult ionization).

Methyl-5-(4-fluoro-2-(prop-2-ynyloxy)phenyl)isoxazole-3-carboxylate (2-43): Off-white solid. Yield = 38%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 8.0233 – 7.99 (m, 1H), 7.27 (dd, *J* = 11.3 Hz, *J* = 2.39 Hz, 1H), 7.15 (s, 1H), 7.09 – 7.03 (m, 1H), 5.12 (d, *J* = 2.31 Hz, 2H), 3.94 (s, 3H), 3.76 (t, *J* = 2.22 Hz, 1H); ESI-MS: *m/z* $[M+H]^+$ = 274.7.

Methyl 5-(5-chloro-2-(prop-2-ynyloxy)phenyl)isoxazole-3-carboxylate (2-49): Yellow solid. Yield = 34%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.97 (d, *J* = 2.79 Hz, 1H), 7.46 – 7.40 (dd, *J* = 9.29 Hz, *J* = 2.61 Hz, 1H), 7.21 (s, 1H), 7.07 (d, *J* = 9.06 Hz, 1H) 4.85 (d, *J* = 2.10 Hz, 2H), 3.99 (s, 3H), 2.59 (t, *J* = 2.10 Hz, 1H); ESI-MS: *m/z* $[M+H]^+$ = 291.8.

5-(2-Methyl-4-prop-2-ynyloxy-phenyl)-isoxazole-3-carboxylic acid methyl ester (2-50): Pale brown solid. Yield = 44 %. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.73-7.70(d, *J* = 8.37 Hz, 1H), 7.07(s, 1H), 7.00-6.97(m, 2H), 4.87-4.86(d, J = 2.2 Hz, 2H), 3.9 (s, 3H), 3.58-3.57 (t, J = 4.44 Hz, 1H), 2.45(s, 3H). ^{13}C-NMR (75 MHz, DMSO-d6) δ 171.12, 160.02, 158.78, 156.26, 138.23, 130.01, 118.95, 117.64, 112.89, 102.29, 78.89, 78.54, 55.57, 52.84, 21.14. ESI-MS: *m/z* = N.A. (didn't show up in both +ve and –ve modes, probably due to difficult ionization).

Methyl 5-(2-(prop-2-ynyloxy)phenyl)isoxazole-3-carboxylate (2-51): Dark brown solid. Yield = 42 %. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.94 (dd, *J* = 7.74 Hz, *J* = 1.65 Hz, 1H), 7.56 – 7.53 (m, 1H), 7.32 (d, *J* = 8.22 Hz, 1H), 7.20 – 7.15 (m, 2H), 5.31 (d, *J* = 2.46 Hz, 2H), 3.93 (s, 3H), 3.66 (t, *J* = 2.46 Hz, 1H); ESI-MS: *m/z* = N.A. (didn't show up in both +ve and –ve modes, probably due to difficult ionization).

General procedure for the hydrolysis of the methyl esters (2-F, 2-G, 2-H, 2-I, 2-J): The methyl ester (4 mmol) was suspended in methanol (10 ml) and a NaOH solution (10 ml; 10 M solution) was added slowly and the reaction was stirred for 3 h before the pH was adjusted to ~2 using HCl (2 N solution) under cold conditions. The resulting precipitate was collected, washed with cold water, dried *in vacuo* to furnish the desired acid. Compounds **2-F, 2-G, 2-H, 2-I, 2-J** were prepared starting from **2-44, 2-45, 2-52, 2-53, 2-54** respectively.

5-(3,5-dibromo-2-(prop-2-ynyloxy)phenyl)isoxazole-3-carboxylic acid (2-F): Brown powder. Yield = 92%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 8.13 (d, *J* = 2.46 Hz, 1H), 8.08 (d, *J* = 2.28 Hz, 1H), 7.31 (s, 1H), 4.82 (d, *J* = 2.28 Hz, 2H), 3.63 (t, *J* = 2.39 Hz, 1H); ^{13}C-NMR (125 MHz, DMSO-d6) δ 165.2, 160.6, 157.9, 150.6, 137.4, 129.7, 124.3, 119.6, 118.1, 105.2, 100.0, 80.0, 77.7, 60.5; ESI-MS: *m/z* $[M-H]^-$ = 400.0.

5-(4-fluoro-2-(prop-2-ynyloxy)phenyl)isoxazole-3-carboxylic acid (2-G): Pale brown solid. Yield = 92%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.95 (t, *J* = 7.65 Hz, 1H), 7.21 (d, *J* = 11.19 Hz, 1H), 7.08 (s, 1H), 7.00 (t, *J* = 8.48 Hz, 1H), 5.08 (s, 2H), 3.71 (s, 1H); ^{13}C-NMR (75 MHz, DMSO-d6) δ 165.8 (d, J_{CF} = 2.18 Hz), 166.2, 162.5, 160.9, 157.7, 155.5 (d, J_{CF} = 10.9 Hz), 129.0 (d, J_{CF} = 10.4 Hz), 112.0 (d, J_{CF} = 3.3 Hz), 108.7 (d, J_{CF} = 21.8 Hz), 103.2, 102.0 (d, J_{CF} = 26.2 Hz), 79.3, 78.1, 56.9; ESI-MS: *m/z* $[M+H]^+$ = 262.1.

5-(3-methoxy-4-(prop-2-ynyloxy)phenyl)isoxazole-3-carboxylic acid (2-H): White powder. Yield = 96%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.53 (br s, 2H), 7.38 (s, 1H), 7.19 (d, *J* = 8.88 Hz, 1H), 4.91 (d, *J* = 1.47 Hz, 2H), 3.90 (s, 3H), 3.62 (s, 1H); ^{13}C-NMR (75 MHz, DMSO-d6) δ 170.9, 161.0, 157.7, 149.5, 148.6, 119.9, 118.6, 113.9, 109.4, 100.0, 78.9, 78.7, 56.1, 55.9; ESI-MS: *m/z* $[M+H]^+$ = 274.7.

5-(2-Methyl-4-prop-2-ynyloxy-phenyl)-isoxazole-3-carboxylic acid (2-I): Brown solid. Yield = 93%.

5-(2-(prop-2-ynyloxy)phenyl)isoxazole-3-carboxylic acid (2-J): Dull white solid. Yield = 92%. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.94 (dd, *J* = 7.88 Hz, *J* = 1.64 Hz, 1H), 7.59 – 7.53 (m, 1H), 7.32 (d, *J* = 8.22 Hz, 1H), 7.19 (t, *J* = 7.56 Hz, 1H) 7.15 (s, 1H), 5.06 (d, *J* = 2.46 Hz, 2H), 3.68 (t, *J* = 2.22 Hz, 1H); ESI-MS: *m/z* $[M-H]^-$ = 242.1.

2.3.3 General Procedure for the High-throughput 'click' assembly

A total of 10 different alkynes (**2-A** to **2-J**) and 325 azides (Azides synthesized by the above solid-phase method as well as few azides synthesized from the solution phase) were subjected to Cu(I) catalyzed 1,3-dipolar cycloaddition which to assemble the 3250 member library. The library was assembled in 14-different 384 well plates.

Synthesis of the 3250 member click library:

The optimized condition for the alkyne-alkyne click reaction is given in the **Table 7.1**. Alkyne solution (25 mM, 10 μL) was dispensed into each well of the 384-well plate using a Multi-drop Combi dispenser (Thermo Scientific) followed by the transfer of azide solutions (50 mM, 7 μL) from the 96-well storage plate to the first quadrant of the 384-well reaction plate using a Sciclone ALH 3000 Liquid Handler Workstation (Caliper Life Sciences). Similarly other azides were transferred from the 96-well storage plates to the rest of the three quadrants of the 384-reation plate. t-BuOH (43 μl) and catalyst mix$^{\pm}$(40 μl) were added to each well in the reaction plate using the bulk liquid dispenser. The 384-well plate was carefully sealed using a silicone based cap-mat and shaken for 2 days. The solvent was dried in vacuo and the resulting residue was redissolved in DMSO (100 μL/well). The 'click' products were analyzed by LC-MS to determine the quality and

identity (See Appendix). Representative members were also characterized by 1H NMR analysis and the data is presented below. The concentration of the product in each well is calculated to be 2.5 mM based on the assumption that all alkynes were reacted quantitatively. The compounds were stored in -20^{o} C.

$^{\pm}$Catalyst mix is a mixture of $CuSO_4$ solution (25 mM, 2 μL) and Sodium ascorbate (50 mM, 5 μL) in water (33 μL).

No.	Component	Equivalent	Concentration (mM)	Volume (μL)
1	Alkyne (in DMSO)	1	25 mM	10
2	Azide (in DMSO)	1.4	50 mM	7
3	$CuSO_4$.5 H2O (in Water)	0.2	25 mM	2
4	Sodium ascorbate (in Water)	0.5	50 mM	5
5	*t*-BuOH	-	-	43
6	Water	-	-	33

Table 7.1 Reaction conditions for click reaction

NMR Characterization of few Final products:

5-(4-fluoro-2-((1-(2-(3-fluorobenzylamino)-2-oxoethyl)-1H-1,2,3-triazol-4-yl)methoxy)phenyl)isoxazole-3-carboxylic acid (2-A5-2C-W2): ^{1}H NMR (300 MHz, DMSO-d6) δ 4.34 (m, 2H), 5.23 (s, 2H), 5.45 (s, 2H), 6.97-7.01 (m,1H), 7.03 (s, 1H), 7.08-7.13 (m, 2H), 7.33-7.41 (m, 1H), 7.44- 7.48 (dd, 1H, J= 11.34Hz, J= 2.31Hz), 7.93-

7.99 (m, 1H), 8. 28(s,1H), 8.88 (t, 1H, J= 5.76Hz); ESI-MS (IT-TOF) m/z = $[M+H]^+$ 470.1288

5-(4-fluoro-2-((1-(2-(3-fluoro-4-methylphenylamino)-2-oxoethyl)-1H-1,2,3-triazol-4-yl)methoxy)phenyl)isoxazole-3-carboxylic acid (2-I-2-A6-2C): ^{1}H NMR (300 MHz, DMSO-d6) δ 2.17 (s, 3H), 5.36 (s, 2H), 5.47 (s, 2H), 6.98-7.01 (m, 1H), 7.03 (s, 1H), 7.17-7.24 (m, 2H), 7.45-7.52 (m, 2H), 7.96 (m, 1H), 8.33 (s, 1H), 10.59 (s, 1H); ESI-MS (IT-TOF) m/z = $[M+H]^+$ 470.1295

5-(4-fluoro-2-((1-(2-(4-(methylthio)phenylamino)-2-oxoethyl)-1H-1,2,3-triazol-4-yl)methoxy)phenyl)isoxazole-3-carboxylic acid (2-I-2-B7-2C): ^{1}H NMR (300 MHz, DMSO-d6) δ 2.45 (s, 3H), 5.35 (s, 2H), 5.46 (s, 2H), 6.98-7.01 (m, 1H), 7.03 (s, 1H), 7.24 (d, 2H, J= 8.70 Hz), 7.49-7.55 (m, 3H), 7.93-7.99 (m, 1H), 8.33 (s, 1H), 10.48 (s, 1H)); ESI-MS (IT-TOF) m/z = $[M+H]^+$ 484.1053

5-(4-fluoro-2-((1-(2-oxo-2-(4-(trifluoromethoxy)phenylamino)ethyl)-1H-1,2,3-triazol-4-yl)methoxy)phenyl)isoxazole-3-carboxylic acid (2-I-2-B8-2C): ^{1}H NMR (300 MHz, MeOH-d4) δ 5.39 (s, 2H), 5.44 (s, 2H), 6.88-6.95 (m, 1H), 7.05 (s, 1H), 7.23-7.25 (m, 2H), 7.28-7.29(m,1H), 7.66-7.69 (m, 2H), 7.95-8.00 (m, 1H), 8.26 (s, 1H)); ESI-MS (IT-TOF) m/z = $[M+H]^+$ 522.1054

5-(4-fluoro-2-((1-(2-(4-methoxyphenylamino)-2-oxoethyl)-1H-1,2,3-triazol-4-yl)methoxy)phenyl)isoxazole-3-carboxylic acid (2-I-2-B9-2C): ^{1}H NMR (300 MHz,

DMSO-d6) δ 3.75 (s, 3H), 5.33 (s, 2H), 5.46 (s, 2H), 6.88-6.91 (d, 2H, J= 8.88 Hz), 7.01-7.09 (m, 2H), 7.45-7.49 (m, 3H), 7.93-7.99 (m, 1H), 8.32 (s, 1H), 10.34 (s, 1H)); ESI-MS (IT-TOF) m/z = $[M+H]^+$ 468.1424

5-(2-((1-(2-(3,4-dimethoxyphenylamino)-2-oxoethyl)-1H-1,2,3-triazol-4-yl)methoxy)-4-fluorophenyl)isoxazole-3-carboxylic acid (2-I-2-B10-2C): ^{1}H NMR (300 MHz, DMSO-d6) δ 3.71 (s, 6H), 5.33 (s, 2H), 5.47 (s, 2H), 6.89-6.91 (d, 1H, J= 8.70 Hz), 7.02-7.05 (m, 3H), 7.29 (s, 1H), 7.45-7.50 (dd, 1H, J= 11.34Hz, J= 2.31 Hz), 7.94-7.97 (m, 1H), 8.32 (s, 1H), 10.35 (s, 1H); ESI-MS (IT-TOF) m/z = $[M+H]^+$ 498.1454

5-(4-fluoro-2-((1-(2-oxo-2-(4-phenoxyphenylamino)ethyl)-1H-1,2,3-triazol-4-yl)methoxy)phenyl)isoxazole-3-carboxylic acid (2-I-2-B11-2C): ^{1}H NMR (300 MHz, DMSO-d6) δ 5.36 (s, 2H), 5.47 (s, 2H), 6.95-7.13 (m, 7H), 7.33-39 (m, 2H), 7.45-7.50 (dd, 1H, J= , J= Hz), 7.54-7.58 (m, 2H), 7.97 (m, 1H), 8.33 (s, 1H), 10.51 (s, 1H); ESI-MS (IT-TOF) m/z = $[M+H]^+$ 530.1579

5-(4-Fluoro-2-{1-[(4-pentyl-phenylcarbamoyl)-methyl]-1H-[1,2,3]triazol-4-ylmethoxy}-phenyl)-isoxazole-3-carboxylic acid (2-I-2-C8-2C): ^{1}H NMR (300 MHz, DMSO-d6) δ 0.82-0.87 (m, 3H), 1.22-1.32 (m, 6H), 1.48-1.58 (m, 2H), 5.34 (s, 2H), 5.4 (s, 2H), 6.98-7.01 (m, 1H), 7.05 (s, 1H), 7.14 (d, 2H, J= 8.37 Hz), 7.45-7.50 (m, 3H), 7.94-7.97 (m, 1H), 8.32 (s, 1H), 10.39 (s, 1H); ESI-MS (IT-TOF) m/z = $[M+H]^+$ 508.2006.

7.3 Solid-Phase Assembly and *in situ* screening of Protein Tyrosine Phosphatases inhibitors

7.3.1 Experimental details of the synthesis of PTP traceless library

Synthesis of linkers:

To a solution of bromoalkynoic acid (15 mmol) in dichloromethane were added benzyl alcohol (15 mmol), DMAP (3 mmol) and EDC (16.5 mmol) at ice cold condition. The solution was stirred for 15 min at room temperature, after which the organic phase was extracted with extracted with $NaHCO_3$ (2 x 50 mL), water (2 x 50 mL) and brine (1 x 30 mL). The organic phase was separated and then dried using anhydrous Na_2SO_4 and the solvent was removed *in vacuo* to afford the crude product, which, upon further purification by flash column chromatography afforded the pure product. **3-9a** and **3-9b** were obtained from 6-bromohexanoic acid and 8-bromooctanoic acid respectively in good yields.

Benzyl-6-bromohexanoate (3-9a): Yield = 89%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.35-7.39 (m, 5H), 5.10 (s, 2H)< 3.38-3.33 (t, *J* = 6.72 Hz , 2H), 2.38-2.33 (t, *J* = 7.4 Hz, 2H), 1.88-1.79 (m, 2H), 1.71-1.60(m, 2H), 1.49-1.39(m, 2H). ^{13}C-NMR (300 MHz, $CDCl_3$) δ 173.14, 135.98, 128.50, 128.15, 66.11, 33.97, 33.41, 32.31, 27.5, 24.01.

Benzyl-8-bromooctanoate (3-9b): Yield = 83%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.33-7.28 (m, 5H), 5.10 (s, 2H), 3.38-3.33 (t, *J* = 6.6 Hz 2H), 2.36-2.31 (t, *J* = 7.6 Hz, 2H),

1.85-1.76 (m, 2H), 1.63-1.42 (m, 2H), 1.39-1.29 (m, 6H). ^{13}C-NMR (300 MHz, $CDCl_3$) δ 173.44, 136.11, 128.14(m), 66.04, 34.18, 33.83, 32.66, 28.84, 28.34, 27.92, 24.78.

Synthesis of Warhead:

1-(4-Benzyloxy-phenyl)-ethanone (3-2): To a solution of 4-hydroxyacetophenone **3-1** (0.11 mol, 15 g) in acetone (150 ml) were added K_2CO_3 (0.22 mol, 30 g) and benzyl bromide (0.12 mol, 14.4 ml). The resulting suspension was refluxed for 12 hrs. After which the leftover K_2CO_3 was removed by filtration. On removal of acetone under vacuum, the residue obtained was dissolved in Dichloromethane (200 ml) and extracted with water (3 x 150 ml) and brine (1 x 100ml). The organic layer was dried over anhydrous sodium sulfate and the solvent was removed under vacuum. The resulting crude solid was suspended in Hexane: Ethyl acetate mixture (9:1) and heated to boiling. On slow cooling shiny white crystals precipitated out which were filtered and washed with hexane and dried to afford **3-2** (20.68 g, 83 %). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.94-7.90 (d, *J* = 9.54 Hz, 2H), 7.44-7.38 (m, 5H), 7.01-6.97(d, *J* = 11.67 Hz, 2H), 5.12 (s, 2H), 2.54 (s, 2H). ^{13}C-NMR (300 MHz, $CDCl_3$) δ 197, 167.9, 136.09, 130.54-130.42(d), 128.62, 128.17, 127.40, 114.46, 70.04-69.91(d), 21.88. ESI-MS: *m/z* $[M+H]^+$ = 227.1

5-(4-Benzyloxy-phenyl)-isoxazole-3-carboxylic acid methyl ester (3-3): To a mixture of benzyl-derivatized acetophenone **3-2** (20g, 80 mmol) and dimethyl oxalate (10.38g, 88 mmol) was added freshly prepared NaOMe (0.5 M in MeOH, 84 mmol). The reaction mixture was refluxed for 12 h and then cooled to room temperature. To the same reaction

flask MeOH (80 ml), $NH_2OH.HCl$ (6.1g, 88 mmol) and a catalytic amount of *p*-$TsOH.H_2O$ were added and the resulting mixture was refluxed continuously for 3 days. Upon cooling to room temperature, the precipitated compound was collected, washed with water and ice-cold methanol to afford the pure isoxazole-3-carboxylic acid methyl ester **3-3** (12.57g, 42% in 2 steps)**.** ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.91-7.87 (d, *J* = 8.85 Hz, 2H), 7.49-7.35 (m, 5H), 7.32 (s, 1H), 7.20-7.17 (d, *J* = 8.88 Hz, 2H), 5.20 (s, 2H), 3.92 (s, 3H). ESI-MS: *m/z* $[M+Na]^+$ = 332.1 and *m/z* $[2M+Na]^+$ = 640.9

5-(4-Benzyloxy-phenyl)-isoxazole-3-carboxylic acid (3-4): To a vigorously stirred suspension of **3-3** (10.8g, 35mmol) in MeOH (100 ml) was added NaOH (100ml, 10 N) and allowed to stir for 4 hours after which MeOH was removed under reduced pressure. To the resulting suspension, water (100 ml) was added and the pH was adjusted to ~ 2 using 2N HCl solution. The white precipitate formed was filtered, washed with ice-cold water and dried to afford the acid **3-4** (9.58g, 93%), which was used without further purification. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.91-7.88 (d, *J* = 8.37 Hz, 2H), 7.49-7.35 (m, 5H), 7.26 (s, 1H), 7.20-7.17 (d, *J* = 1.4 Hz, 2H), 5.2 (s, 2H). ^{13}C-NMR (300 MHz, DMSO-d6) δ 171.18, 161.30, 160.61, 158.06, 136.92, 128.85-127.91(m), 119.46, 99.76, 69.81. ESI-MS: *m/z* $[M-H]^-$ = 294.0

5-(4-Benzyloxy-phenyl)-isoxazole-3-carboxylic acid tert-butyl ester (3-5): To the solution of the acid **3-4** (9g, 30.5 mmol) in THF (100 ml) were added t-BuOH (6.7 g, 91.5 mmol), EDC (8.74 g, 45.75 mmol) and DMAP (1.86 g, 15.25 mmol) at 0^{o}C. After a few minutes a cloudy suspension was formed, on adding DCM (40 ml) to the suspension

a clear solution was formed. The resulting clear solution was allowed to stir for 12 hours at room temperature. The solvents were removed under reduced pressure and the resulting crude mass was directly purified using flash chromatography to afford **3-5** (8.24g, 77%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.73-7.70 (d, *J* =10.89 Hz, 2H), 7.44-7.33 (m, 5H), 7.06-7.03 (d, *J* = 8.7 Hz, 2H), 6.72 (s, 1H), 5.11 (s, 2H), 1.63 (s, 9H). ^{13}C-NMR (300 MHz, $CDCl_3$) δ 171.33, 160.59,159.21, 158.06, 136.28, 128.66-127.44(m), 119.81, 115.38, 98.59, 83.48, 70.12, 28.06. ESI-MS: *m/z* $[M+Na]^+$ = 374.0 and *m/z* $[2M+Na]^+$ = 724.9.

5-(4-Hydroxy-phenyl)-isoxazole-3-carboxylic acid tert-butyl ester (3-6): To a solution of benzyl ether **3-5** (8g, 23mmol), in methanol-THF (1:1) was slowly added Pd/C power (0.8g, 10 % wt). The reaction was stirred for 4 h under hydrogen atmosphere. After which, Pd/C was filtered out and the solvent was removed completely under reduced pressure to yield pure derivatized phenol **3-6** in quantitative yield. ^{1}H-NMR (300 MHz, DMSO-d6) δ 7.78-7.54 (d, *J* = 8.7 Hz, 2H), 7.14 (s, 1H), 6.93-6.90 (d, *J* = 8.7 Hz, 1.57 (s, 9H). ^{13}C-NMR (300 MHz, $CDCl_3$) δ 171.74, 160.39, 159.05, 158.05, 128.06, 117.63, 116.40, 98.80, 83.43, 28.01 ESI-MS: *m/z* $[M+Na]^+$ = 283.9 and *m/z* $[2M+Na]^+$ = 545.1.

General procedure for the synthesis of the phenyl ethers (3-7a and 3-7b): Potassium carbonate (22 mmol) and catalytic amount of dibenzo-18-crown-6 were added to a solution of bromo-benzylalkynolate (2.5mmol) and derivatized phenol **3-6** (11 mmol) in acetonitrile (100 ml) followed by refluxing for 3 h, after which the organic phase was removed under reduced pressure and taken into dichloromethane layer (80 ml) and

extracted with water (2 x 40 ml) and brine (1 x 40 ml). The organic phase was then dried with anhydrous Na_2SO_4, and the solvent was removed *in vacuo* to afford the crude product, which, upon further purification by flash column chromatography, afforded the pure ether. Compounds **3-7a** and **3-7b** were prepared from compound **3-9a** and **3-9b** respectively on treatment with compound **3-6**.

5-[4-(5-Benzyloxycarbonyl-pentyloxy)-phenyl]-isoxazole-3-carboxylic acid tert-butyl ester (3-7a): Yield = 94%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.72-7.69 (d, *J* = 8.7 Hz, 2H), 7.35 (m, 5H), 6.96-6.93 (d, *J* = 8.88 Hz, 2H), 6.71 (s, 1H), 5.12 (s, 2H), 4.01-3.97 (t, *J* = 6.04 Hz, 2H), 2.43-2.38 (t, *J* = 7.48 Hz, 2H), 1.90-1.79 (m, 4H), 1.63 (s, 9H), 1.63-1.43 (m, 2H). ^{13}C-NMR (300 MHz, $CDCl_3$) δ 173.28, 171.39, 160.86, 159.18, 157.99, 135.96, 128.48-128.13(m), 127.44, 119.39, 114.90, 98.40, 83.41, 67.76, 66.10, 34.09, 28.71, 28.01, 27.54, 25.51, 24.46, 23.99. ESI-MS: *m/z* $[M+H]^+$ = 465.8, $[M+Na]^+$ = 488.1 and *m/z* $[2M+Na]^+$ = 953.0

5-[4-(7-Benzyloxycarbonyl-heptyloxy)-phenyl]-isoxazole-3-carboxylic acid tert-butyl ester (3-7b): Yield = 63%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.73-7.70 (d, *J* = 8.7 Hz, 2H), 7.35 (m, 5H), 6.97-6.95 (d, *J* = 8.7 Hz, 2H), 6.72 (s, 1H), 5.12 (s, 2H), 4.01-3.97 (t, *J* = 6.43 Hz, 2H), 2.39-2.34 (t, *J* = 7.48 Hz, 2H), 1.83-1.66 (m, 4H), 1.63(s, 9H), 1.47-1.37(m, 6H).

General procedure for the benzyl ester hydrogenolysis (3-A and 3-B):

To a solution of benzyl ester (5g, 11mmol) in THF, was slowly added Pd/C (0.5g, 10 % wt). The reaction was stirred for 4 h under hydrogen atmosphere. After which, Pd/C was filtered out and the solvent was removed completely under reduced pressure to yield pure the acid derivative in quantitative yield. Hydrogenolysis of compounds **3-7a** and **3-7b** gave the products **3-A** and **3-B** respectively.

5-[4-(5-Carboxy-pentyloxy)-phenyl]-isoxazole-3-carboxylic acid tert-butyl ester (3-A): Yield = 98%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.73-7.70 (d, *J* = 9.03 Hz, 2H), 6.98-6.95 (d, *J* = 8.7 Hz, 2H), 6.73 (s, 1H), 4.04-3.99 (t, *J* = 6.27 Hz, 2H), 2.43-2.38 (t, *J* = 7.48 Hz, 2H), 1.93-1.68 (m, 4H), 1.63 (s, 9H), 1.59-1.47 (m, 2H). ^{13}C-NMR (300 MHz, $CDCl_3$) δ 179.61, 171.43, 160.90, 159.21, 158.01, 127.28, 119.44, 114.95, 98.45, 83.48, 67.78, 33.87, 28.74, 28.03, 25.49, 24.33. ESI-MS: *m/z* $[M-1]^-$= 374.1 and *m/z* $[2M-1]^-$ = 749.0.

5-[4-(7-Carboxy-heptyloxy)-phenyl]-isoxazole-3-carboxylic acid tert-butyl ester (3-B): Yield = 98%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.73-7.7 (d, *J* = 8.88 Hz, 2H), 7.26(s, 1H), 6.98-6.95 (d, *J* = 8.88 Hz, 2H), 4.02-3.98 (t, *J* = 6.42 Hz, 2H), 2.39-2.34 (t, *J* = 7.39 Hz, 2H), 1.83-1.76 (m, 2H), 1.66 (m, 2H), 1.63 (s, 9H), 1.52-1.39 (m, 6H). ^{13}C-NMR (300 MHz, $CDCl_3$) δ 179.90, 172.19, 161.72, 161.72, 159.97, 158.75, 128.21, 120.9, 115.67, 99.15, 84.20, 68.80, 34.57, 29.73, 29.64, 28.77, 26.49, 25.27. ESI-MS: *m/z* $[M-H]^-$ = 402.2 and *m/z* $[2M-1]^-$ = 805.1.

General procedure for the activation of acid to acid chloride (3-8a and 3-8b):

A solution of the carboxylic acid (6.6mmol) in freshly distilled dichloromethane (25 ml) was added DIEA (13.2mmol) followed by the slow addition of fresh oxalyl chloride (6.6 mmol) using an air-tight syringe at 0^{o}C. After the completion of the addition, the reaction mixture was allowed to stir at room temperature for another 4 h. The completion of the reaction was monitored by TLC. The reaction mixture was used as such for further solid-phase coupling reaction without any further purification. The acid chlorides **3-8a** and **3-8b** were prepared from acids **3-A** and **3-B** respectively.

Solid-phase assembly of the PTP inhibitors

The construction of the PTP inhibitor library was achieved by IRORI® split-and-pool directed sorting technology. Final products were released from the solid phase by our optimized TFA cleavage protocol.

Synthesis of the reductive aminated resin (3-10):

PL-FMP resin (70 x ~30 mg, 0.9 mmol/g) was taken in 70 MacroKan™ reactors each containing a RF tag. The resin was swelled in 1,2-dichloroethane (200 mL) for about two hours after which the solvent was decanted. The 70 microreactors were then distributed equally into 35 different bottles of capacity 15 mL containing 2% acetic acid in 1,2-dichloroethane (6 mL). Amines (**3-a1** to **3-a35**) (5 eq) were added to the bottles, so that each bottle contains a unique amine. The reaction mixture was then incubated for about 3 hours after which sodium triacetoxyborohydride (6 eq) was added. After shaken for another 8 hours the solution was decanted and the reactors were combined and washed

with DCM (150 mL x 5), MeOH (150 mL x 2) and THF (200 mL x 3) and dried to afford resin **3-10**.

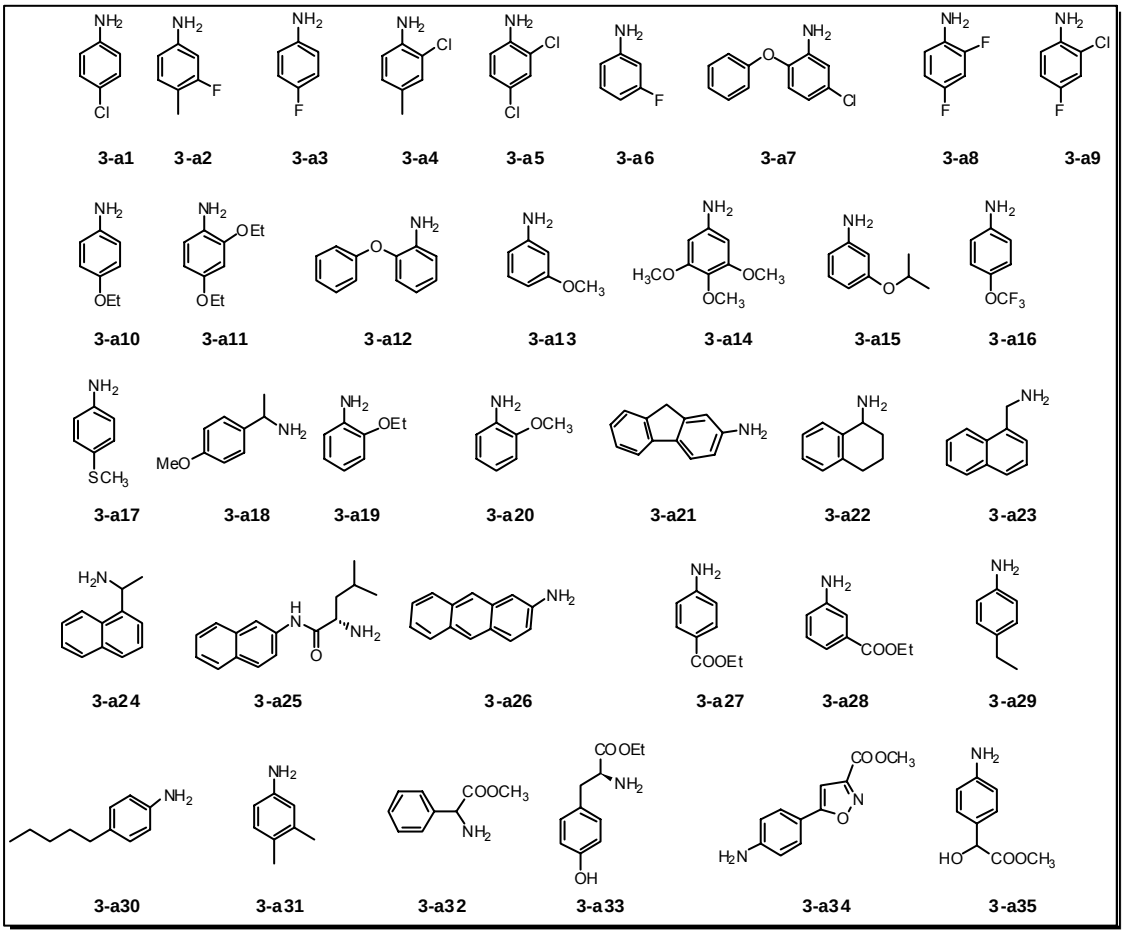


Fig. 7.7 List of 35 different amines used for the reductive amination

Synthesis of the N-acylated resin (3-11a and 3-11b):

The 70 reductive aminated resins were separated into 2 sets (35 different reductive aminated resins x 2 sets). To the first set DCM (50 mL) and DIEA (5 eq) were added and cooled using an ice-bath. The reaction mixture consisting of acid chloride, **3-8a** (6.6 mmol) was then slowly added to the bottle under a stream of argon. The reaction mixture was shaken for 12 h at room temperature after which the solution was decanted and the

resin was washed with DCM (100 ml x 3), MeOH (100 mL x 2) and THF (100 mL x 3) and dried to afford resin **3-11a**. Similarly the second set of resin on acylation with **3-8b** gave the resin **3-11b**.

Protocol for the cleavage and release of the inhibitors from the solid-support (3-A1 to **3-A35 and 3-B1** to **3-B35):**

Each dried resin was treated with a 1.3 mL solution containing TFA (50%), DCM (45 %) and water (5%) and the mixture was shaken for 3.5 hours and transferred into a 96-well plates and concentrated in vacuo to afford the compounds **3-A1** to **3-A35** and **3-B1** to **3-B25**. Representative yield varied from 50% to 90%. The compounds were then redissolved in DMSO. Most of the compounds synthesized from the above protocols are of high purity (90-95%). Representative compounds were further characterized without any purification by LCMS and ^{1}H NMR.

NMR Characterization of few Final products:

Compound 3-A1: ^{1}H-NMR (300 MHz, DMSO-d6) δ 10.02 (s, 1H), 7.87-7.86 (d, *J* = 8.7 Hz, 2H), 7.64-7.60 (d, *J* = 8.88 Hz, 2H), 7.35-7.32 (d, *J* = 8.88 Hz, 2H), 7.24 (s, 1H), 7.10-7.07 (d, J = 8.7 Hz, 2H), 4.08-4.04 (t, *J* = 6.24 Hz, 2H), 2.37-2.32 (t, *J* = 7.23 Hz, 2H), 1.87-1.72 (m, 2H), 1.69-1.59(m, 2H), 1.51-1.41 (m, 2H).

Compound 3-A2: ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.99 (s, 1H), 7.87-7.84 (d, *J* = 8.7 Hz, 2H), 7.56-7.52 (d, *J* = 12.84 Hz, 1H), 7.23 (s, 1H), 7.18-7.16 (m, 2H), 7.09-7.06 (d, *J*

= 8.88 Hz, 2H), 4.08-4.04 (t, *J* = 6.39 Hz, 2H), 2.35-2.30 (t, *J* = 7.23 Hz, 2H), 2.16 (s, 3H), 1.76 (m, 2H), 1.68-1.61 (m, 2H), 1.48-1.43 (m, 2H).

Compound 3-A3: ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.94 (s, 1H), 7.87-7.84 (d, *J* = 8.7 Hz, 2H), 7.63-7.58 (m, 2H), 7.22 (s, 1H), 7.15-7.07 (m, 4H), 4.08-4.04 (t, *J* = 6.42 Hz, 2H), 2.35-2.27 (t, *J* = 7.42 Hz, 2H), 1.77 (m, 2H), 1.66 (m, 2H), 1.51-1.41 (m, 2H).

Compound 3-A12: ^{1}H-NMR (300 MHz, $CDCl_3$) δ 9.95 (s, 1H), 8.42-8.39(d, *J* = 8.01 Hz, 1H), 7.81(m, 1H), 7.69-7.66(d, *J* = 7.65 Hz, 1H), 7.37-7.32(m, 2H),7.15-7.10(m, 2H), 7.02-6.99(m, 3H), 6.95-6.92(d, *J* = 8.34 Hz, 1H), 6.87-6.84(d, *J* = 7.65 Hz, 1H), 6.7(m, 1H), 3.97(m, 2H), 2.47-2.39(m, 2H), 1.81-1.70(m, 4H), 1.67-1.53 (m, 2H), 0.88-0.83(m, 2H).

Compound 3-A16: ^{1}H-NMR (300 MHz, DMSO-d6) δ 10.09 (s, 1H), 7.87-7.84(d, *J* = 8.55 Hz, 2H), 7.71-7.84 (d, *J* = 8.88 Hz, 2H), 7.31-7.28 (d, *J* = 8.85 HZ, 2H), 7.24 (s, 1H), 7.10-7.07 (d, *J* = 8.7 Hz, 2H), 4.08-4.04 (t, *J* = 6.24 Hz, 2H), 2.38-2.32 (t, *J* = 7.41 Hz, 2H), 1.77 (m, 2H), 1.64 (m, 2H), 1.51-1.42 (m, 2H).

Compound 3-A17: ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.74-7.69(m, 1H), 7.47-7.44(m, 2H), 7.24-7.22 (m, 3H), 6.97-6.91(m, 2H). 6.79-6.78(m, 1H), 4.39(m, 2H), 3.47(t, *J* = 6.26 Hz, 2H), 2.61(s, 3H), 2.30(m, 2H), 1.70(m, 2H).

Compound 3-A23: ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.99-7.96 (d, *J* = 7.65 Hz, 1H), 7.89-7.83(m, 2H), 7.65-7.63(d, *J* = 8.37 Hz, 2H), 7.57-7.50(m, 2H), 7.44-7.42(m, 2H), 6.91-6.88(d, *J* = 8.7 Hz, 2H), 6.73(s, 1H), 4.93-4.91(d, *J* = 5.22 Hz, 2H), 3.98-3.94(t, *J* = 6.27 Hz, 2H), 2.35-2.27(m, 2H), 1.84-1.68(m, 2H), 1.56-1.45(m, 2H), 0.88-0.83(m, 2H).

Compound 3-A27: ^{1}H-NMR (300 MHz, DMSO-d6) δ 10.21(s, 1H), 7.91-7.84 (m, 4H), 7.74-7.71(d, *J* = 8.7 Hz, 2H), 7.24 (s, 1H), 7.09-7.06 (d, *J* = 8.88 Hz, 2H), 4.31-4.24(q, 2H), 4.08-4.02(t, *J* = 6.24 Hz, 2H), 2.41-2.36(t, *J* = 7.23, 2H), 1.75(m, 2H), 1.65 (m, 2H), 1.52-1.43(m, 2H), 1.33-1.26(t, *J* = 7.05 Hz, 3H).

Compound 3-A30: ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.78(s, 1H), 7.87-7.84(d, *J* = 8.73 Hz, 2H), 7.49-7.46 (d, *J* = 8.37 Hz, 2H), 7.24 (s, 1H), 7.09-7.07 (d, *J* = 8.4 Hz, 4H), 4.08-4.04(t, *J* = 6.42 Hz, 2H), 2.34-2.29 (t, *J* = 7.23 Hz, 2H), 1.77 (m, 2H), 1.66 (m, 2H), 1.57-1.41(m, 4H), 1.32-1.18 (m, 6H), 0.87-0.83(t, *J* = 6.9 Hz, 3H).

Compound 3-A31: ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.70 (s, 1H), 7.87-7.84 (d, *J* = 8.73 Hz, 2H), 7.35-7.28(m, 2H), 7.24 (s, 1H), 7.09-7.06 (d, *J* = 9.06 Hz, 2H), 7.03-7.01(d, *J* = 8.22 Hz, 2H), 4.07-4.03(t, *J* = 6.43 Hz, 2H), 2.33-2.23 (t, *J* = 7.22 Hz, 2H), 2.17-2.15(m, 6H), 1.76 (m, 2H), 1.65 (m, 2H), 1.50-1.40 (m, 2H).

Compound 3-B4: ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.85 (s, 1H), 7.30 (m, 1H), 7.24 (s, 1H), 7.21-7.07 (m, 3H), 6.61-6.49 (d, *J* = 6.9 Hz, 1H), 4.06-4.02(t, *J* = 6.42 Hz, 2H), 3.70 (s, 3H), 2.32-2.27(t, *J* = 7.2 Hz, 2H), 1.73 (m, 2H), 1.59 (m, 2H), 1.35 (m, 6H).

Compound 3-B6: ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.92 (s, 1H), 7.87-7.84(d, *J* = 8.73 Hz, 2H), 7.62-7.58(m, 2H), 7.23(s, 1H), 7.15-7.07(m, 4H), 4.06-4.02(t, *J* = 6.43 Hz, 2H), 2.32-2.27(t, *J* = 7.22 Hz, 2H), 1.74(m, 2H), 1.60(m, 2H), 1.36(m, 6H).

Compound 3-B9: ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.68(s, 1H), 7.87-7.84(d, *J* = 8.7 Hz, 2H), 7.35 (m, 1H), 7.31-7.28(m, 1H), 7.24(s, 1H), 7.09-7.06(d, *J* = 8.88 Hz, 2H), 7.02-7.00 (d, *J* = 8.22 Hz, 1H), 4.06-4.02(t, *J* = 6.42 Hz, 2H), 2.26-2.22(t, *J* = 7.23 Hz, 2H), 2.16-2.14(m, 6H), 1.73(m, 2H), 1.59(m, 2H), 1.35(m, 6H).

Compound 3-B10: ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.70 (s, 1H), 7.87-7.84 (d, *J* = 8.7 Hz, 2H), 7.48-7.45(d, *J* = 9.03 Hz, 2H), 7.23 (s, 1H), 7.09-7.06(d, *J* = 8.88 Hz, 2H), 6.85-6.81(d, *J* = 9.03 Hz, 2H), 4.06-4.01(t, *J* = 6.45Hz, 2H), 3.99-3.92(q, 2H), 2.29-2.24(t, *J* = 7.23 Hz, 2H), 1.73(m, 2H), 1.59(m, 2H), 1.40-1.23(m, 9H).

Compound 3-B16: ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.87 (s, 1H), 7.87-7.84(d, *J* = 8.73 Hz, 2H), 7.56-7.54(d, *J* = 8.7 Hz, 2H), 7.23-7.19(m, 3H), 7.09-7.06(d, *J* = 8.88 Hz, 2H), 4.06-4.02(d, *J* = 6.42 Hz, 2H), 2.32 (m, 2H), 1.73(m, 2H), 1.62(m, 2H), 1.35(m, 6H).

Compound 3-B17: ^{1}H-NMR (300 MHz, DMSO-d6) δ 10.07(s, 1H), 7.86-7.83 (d, *J* = 8.7 Hz, 2H), 7.70-7.67 (d, *J* = 9.06 Hz, 2H), 7.30-7.27 (d, *J* = 8.4 Hz, 2H), 7.20 (s, 1H), 7.09-7.06 (d, *J* = 8.85Hz, 2H), 4.06-4.02(t, *J* = 6.49 Hz, 2H), 2.60 (s, 3H), 2.32-2.27(t, *J* = 7.41 Hz, 2H), 1.73 (m, 2H), 1.61 (m, 2H), 1.36 (m, 6H).

Compound 3-B21: ^{1}H-NMR (300 MHz, DMSO-d6) δ 9.96(s, 1H), 7.93 (s, 1H), 7.87-7.84 (d, *J* = 8.88 Hz, 2H), 7.79-7.74(m, 2H), 7.54-7.51(m, 2H), 7.15 (m, 1H), 7.09-7.06(d, J = 8.88Hz, 2H), 6.99-6.96(m, 1H), 4.06-4.02(t, *J* = 6.40 Hz, 2H), 3.92-3.87(m, 2H), 2.36-3.31(m, 2H), 1.74(m, 2H), 1.63(m, 2H), 1.38(m, 6H).

Compound 3-B23: ^{1}H-NMR (300 MHz, DMSO-d6) δ 8.32 (m, 1H), 8.07-8.46(m, 1H), 7.95-7.92 (m, 1H), 7.87-7.83(m, 3H), 7.57-7.52(m, 2H), 7.49-7.42 (m, 2H), 7.24 (s, 1H), 7.09-7.06 (d, *J* = 8.7 Hz, 2H), 4.73-4.71 (d, *J* = 5.58 Hz, 2H), 4.04-4.00(t, *J* = 6.42 Hz, 2H), 2.18-2.13 (t, *J* = 7.24 Hz, 2H), 1.69 (m, 2H), 1.55 (m, 2H), 1.28 (m, 6H).

7.3.2 Protocol for the Biological Screening:

The activity of phosphatases was assayed by measuring the rate of hydrolysis of the fluorogenic substrate, 6, 8-difluoro-methylumbelliferyl phosphate (DIFMUP) in 25 μl reaction volumes in black polypropylene flat-bottom 384-well microtiter plates (Greiner, Germany). For IC_{50} studies of the 70-member library against PTP1B, dose-dependent reactions were set up by varying the concentrations of the inhibitor, under the same enzyme and substrate concentration. Briefly, a two-fold dilution series of an inhibitor, from approximately 400 μM to 3.125 μM (final concentrations) was prepared. The reaction condition is as shown below.

PTB 1B (2.4 μg/ml) = 5 μl

DIFMUP (20 μM) = 5 μl

Inhibitor (Varied) = 5 μl

Assay Buffer (1X) = 10 μl

[Assay buffer = 25 mM HEPES, 150mM NaCl and 0.1 mg/ml BSA, pH = 7.5]. Negative controls were performed in the absence of enzyme and positive controls were carried out in the presence of enzyme but without inhibitor.

The enzymatic reactions were allowed to incubate at room temperature for 30 min before being initiated by the addition of DIFMUP. The enzymatic reactions were immediately monitored with a Tecan F200 fluorescence plate reader (Tecan, Germany), at λ_{ex} = 355 nm and λ_{em} = 460 nm for a period of 15 min. The IC_{50} was calculated by fitting the fluorescence outputs obtained using the Graphpad Prism software v.4.03 (GraphPad, San Diego) From the preliminary screening results, six representative compounds were identified and their IC_{50} values against PTP1B were again investigated independently. Each IC_{50} plots were generated by averaging the duplicates from two independent assays and results are shown in **Fig. 7.8**.

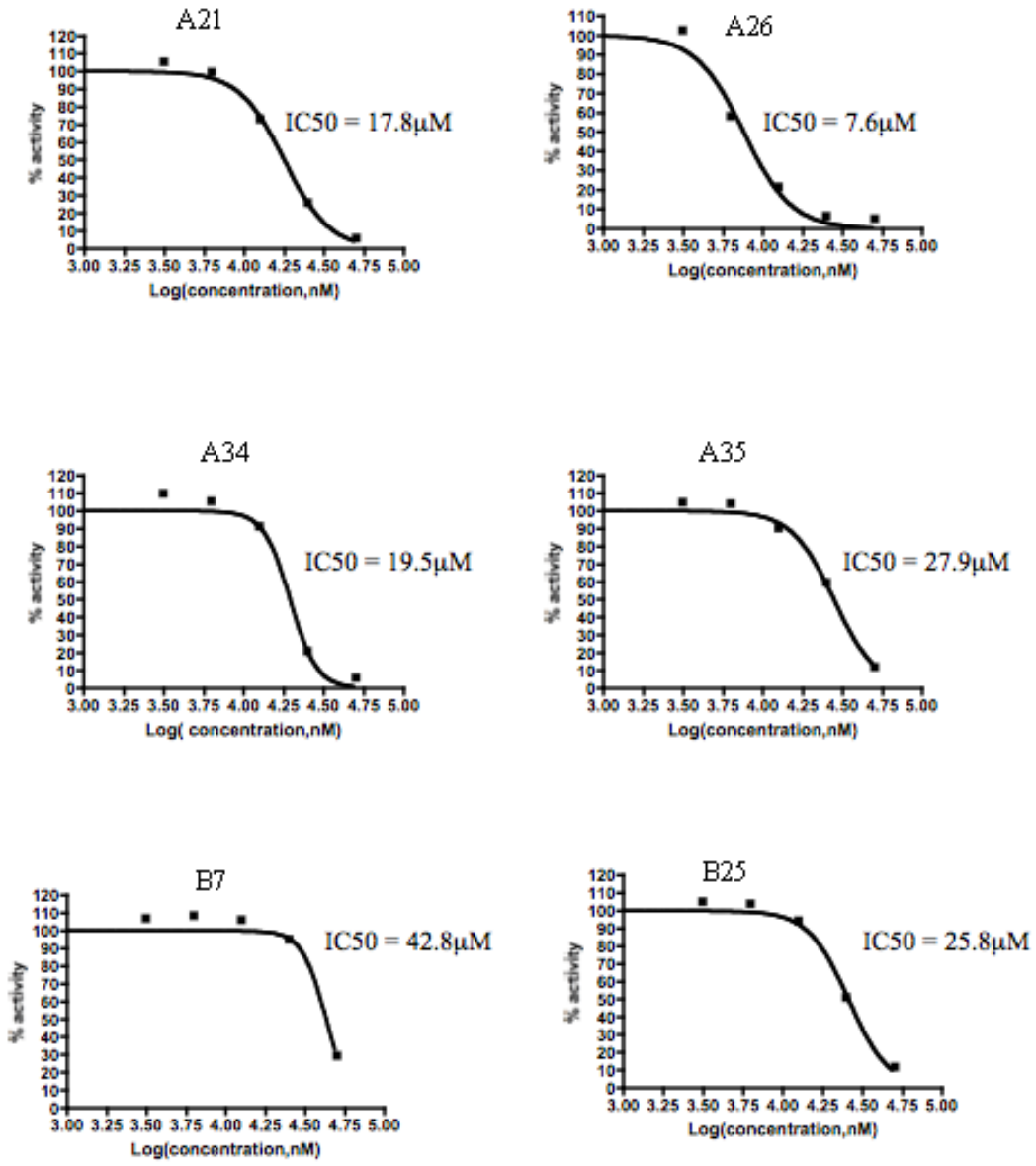


Fig. 7.8 Graphs for determining IC_{50} values of selected library members against PTP1B. The selected members were assayed in duplicate and the results were averaged from two independent experiments followed by obtaining the plots with Graphpad Prism software v4.03 (Graphpad, San Diego).

Selectivity profiles of selected PTP1B inhibitors

For selectivity assays, a panel of PTP enzymes from different PTP-subfamilies was chosen. These include TCPTP, YopH and LMWPTP. The IC_{50} values of the selected library members were determined as previously described for PTP1B. Dose-dependent

reactions were set up by varying the concentrations of the inhibitors; from approximately 100 μM to 3.125 μM (final concentrations). The optimized assay conditions (amount of enzyme needed to reach a fluorescence output of ~30000 in 15 min) of the different phosphates is given below

TCPTP
Enzyme (2.4 μg/ml) = 5 μl
DIFMUP (20 μM) = 5 μl
Inhibitor (Varied) = 5 μl
Assay Buffer (1X) = 10 μl
[Assay buffer = 25 mM HEPES, 150mM NaCl and 0.1 mg/ml BSA, pH = 7.5].

YopH
Enzyme (5.0 μg/ml) = 5 μl
DIFMUP (20 μM) = 5 μl
Inhibitor (Varied) = 5 μl
Assay Buffer (1X) = 10 μl
[Assay buffer = 25 mM HEPES, 150mM NaCl and 0.1 mg/ml BSA, pH = 7.5].

LMWPTP
Enzyme (20 μg/ml) = 5 μl
DIFMUP (20 μM) = 5 μl
Inhibitor (Varied) = 5 μl
Assay Buffer (1X) = 10 μl
[Assay buffer = 50 mM Sodium citrate, 3mM DTT, pH = 6.0].

7.4 Versatile Microwave-Assisted Strategies for the Synthesis of Azide fragments

7.4.1 Experimental details of the synthesis of the azide libraries

General procedure for the solid-phase synthesis of azides (4-1 to 4-10) from PS-TsCl resin:

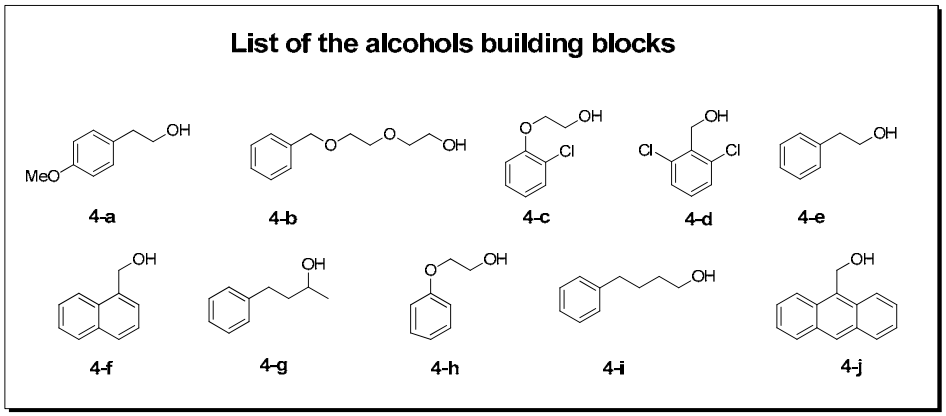


Fig. 7.9 List of alcohol blocks used in route A

To PS-TsCl resin (0.134 mmol) swelled in Dichloromethane (DCM) (3 ml) was added the alcohol (**4-a** to **4-j**) (**Fig. 7.9**) (0.7 mmol) and Pyridine (1.5 ml). The reaction mixture was shaken for 12 hours, after which the reaction solvents were drained and the resin was washed with DCM (3 x 25 ml), DMF (3 x 25 ml), DMF/H_2O (1:1, 2 x 25 ml), THF (3 x 25 ml), DCM (2 x 25 ml) and dried under vacuum to afford the tosylated resin (**4-1a** to **4-10a**). The dried tosylated resin (**4-1a** to **4-10a**) was transferred into microwave reaction vessel followed by the addition of DMF (2 ml) and NaN_3 (1.34 mmol). The resin was microwave irradiated for 30 min after which the resin and the solid residue were filtered off, the DMF layer was taken in to DCM/Ether (20 ml). The organic layer was extracted with water (5 x 10 ml) and brine (1 x 10 ml) dried with anhydrous Na_2SO_4 and concentrated under reduced pressure to afford the azides (**4-1** to **4-10**, 0-92% yield) (**Fig 7.10**)

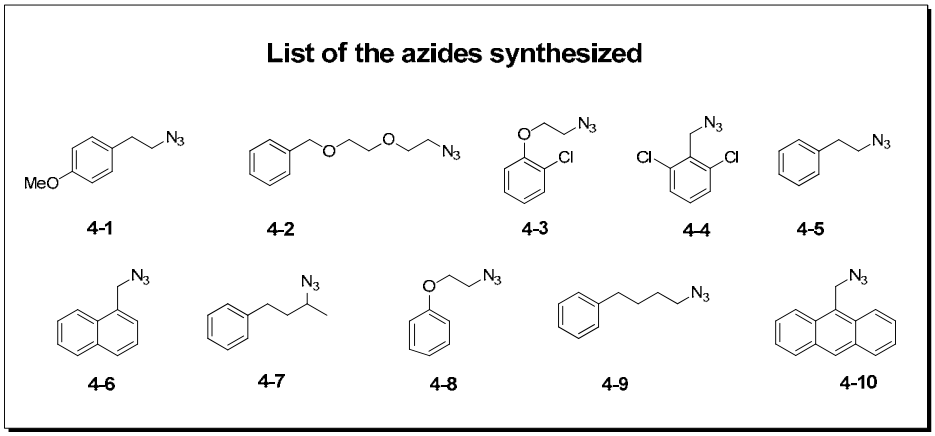


Fig. 7.10 List of azides synthesized from PS-TsCl resin

1-(2-Azido-ethyl)-4-methoxy-benzene (4-1): Yield = 64 %. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.14-7.13 (d, J = 5.31Hz, 2H), 6.89-6.85 (d, J = 5.31 Hz, 2H), 3.79 (s, 3H), 3.48-3.45(t, J = 4.35 Hz, 2H), 2.85-2.82(t, J = 4.35 Hz, 2H). ^{13}C-NMR: (75 MHz, $CDCl_3$) δ 158.50, 130.06, 129.72, 114.09, 55.26, 52.71, 34.50.

((2-(2-azidoethoxy)ethoxy)methyl)benzene (4-2): Yield = 80%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.35 – 7.28 (m, 5H), 4.58 (s, 2H), 3.70 – 3.64 (m, 6H), 3.40 (t, *J* = 5.09 Hz, 2H). ^{13}C-NMR: (75 MHz, $CDCl_3$) δ 138.18, 128.35, 127.68, 127.58, 73.31, 70.73, 70.03, 69.50, 50.72.

1-(2-azidoethoxy)-2-chlorobenzene (4-3): Yield = 70%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.40 – 7.37 (m, 1H), 7.26 – 7.19 (m, 1H), 6.98 – 6.91 (m, 2H), 4.19 (t, *J* = 5.10 Hz, 2H), 3.64 (t, *J* = 5.01 Hz, 2H). ^{13}C-NMR: (500 MHz, $CDCl_3$) δ 153.96, 130.55, 127.70, 123.39, 122.23, 113.76, 68.10, 30.33.

2-(azidomethyl)-1,3-dichlorobenzene (4-4): Yield = 20%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.37 – 7.34 (m, 2H), 7.25 – 7.19 (m, 1H), 4.67 (s, 2H).

(2-azidoethyl)benzene (4-5): Yield = 89%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.35 – 7.21 (m, 5H), 3.50 (t, *J* = 7.23 Hz, 2H), 2.90 (t, *J* = 7.23 Hz, 2H). ^{13}C-NMR: (75 MHz, $CDCl_3$) δ 138, 128.62, 128.41, 126.51, 52.21, 35.07.

1-Azidomethyl-naphthalene (4-6): Yield = 10%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.05 – 7.46 (m, 7H), 4.78 (s, 2H).

(3-Azido-butyl)-benzene (4-7): Yield = 0%. (No Reaction)

(2-azidoethoxy)benzene (4-8): Yield = 92%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.32 – 7.27 (m, 2H), 7.01 – 6.91 (m, 3H), 4.15 (t, *J* = 5.01 Hz, 2H), 3.59 (t, *J* = 5.01 Hz, 2H). ^{13}C-NMR: (500 MHz, $CDCl_3$) δ 158.24, 129.55, 121.38, 114.67, 66.89, 30.32.

(4-Azido-butyl)-benzene (4-9): Yield = 61%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.31 – 7.16 (m, 5H), 3.28 (t, *J* = 6.50 Hz, 2H), 2.64 (t, *J* = 7.23 Hz, 2H), 1.76 – 1.59 (m, 4H). ^{13}C-NMR: (300 MHz, $CDCl_3$) δ 128.41, 125.94, 60.40, 35.40, 30.91, 21.05.

9-Azidomethylanthracene (4-10): Yield = 60 %. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.34-8.28(m, 3H), 8.07-8.04(m, 2H), 7.82-7.79(m, 2H), 7.63-7.49(m, 2H).^{13}C-NMR: (75 MHz, $CDCl_3$) δ 134.83, 130.0, 127.93, 127.55, 125.91, 124.22, 30.39.

MW-assisted solution-phase approaches

General procedure for the conversion of alcohols (4-a, 4-b, 4-d, 4-g and 4-k to 4-p) in to tosylates (4-1b to 4-10b)

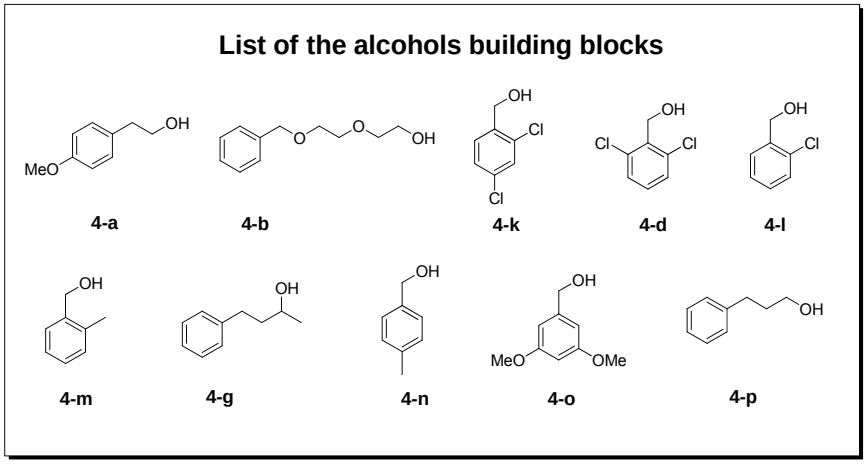


Fig. 7.11 List of alcohol blocks used in tosylation

Commercially available alcohol (**4-a**, **4-b**, **4-d**, **4-g**, **4-k** to **4-p**) (2 mmol) (**Fig. 7.11**) and p-toluenesulfonyl chloride (2.1 mmol) were dissolved in anhydrous diethyl ether (10 ml) and cooled to -5 ºC. Finely powdered KOH (20 mmol) was then added slowly in small portions over a 15-min period. The reaction mixture was further stirred for 4 h, after which the reaction mixture was quenched with water (10 ml) in an ice-bath. The organic layer was separated and the aqueous layer was extracted twice with ether (2 x 15 ml). The combined organic extracts were washed with brine (1 x 15 ml) and dried over Na_2SO_4. The solvent was removed under reduced pressure to afford spectroscopically pure products (**4-1b to 4-10b**) (76 – 99% yield) (**Fig 7.12**).

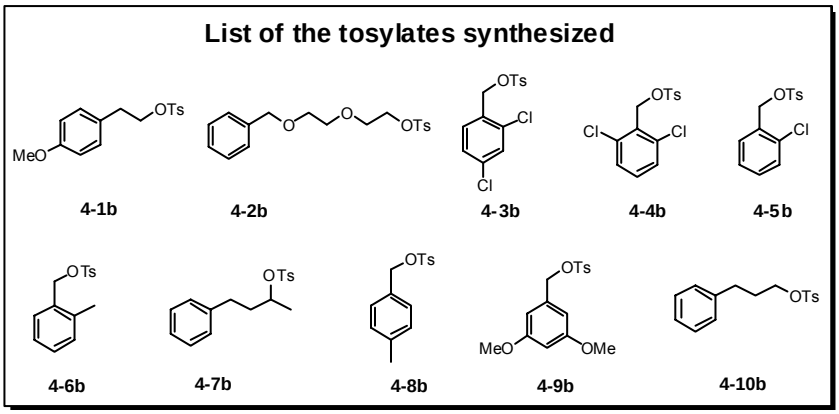


Fig. 7.12 List of tosylates synthesized

2-(4-methoxyphenyl)ethyl 4-methylbenzenesulfonate (4-1b): Yield = 89%.

2-(2-benzyloxyethoxy)ethyl 4-methylbenzenesulfonate (4-2b): Yield = 96%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.80-7.78 (d, *J* = 8.01 Hz, 2H), 7.32-7.29(m, 7H), 4.53(s, 2H), 4.19-4.15(t, *J* = 4.87 Hz, 2H), 3.71-3.57(m, 6H), 2.42(s, 3H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 145.41, 138.75, 133.62, 130.44, 129.02, 128.59, 128.35, 128.28, 73.89, 71.43, 69.98, 69.90, 69.31, 22.24.

2,4-dichlorobenzyl 4-methylbenzenesulfonate (4-3b): Yield = 96%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.80 (d, *J* = 8.22 Hz, 2H), 7.35 – 7.31 (m, 4H), 7.23 – 7.20 (m, 1H), 5.12 (s, 2H), 2.45 (s, 3H).

2,6-dichlorobenzyl 4-methylbenzenesulfonate (4-4b): Yield = 88%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.85 (d, *J* = 8.22 Hz, 2H), 7.34 (d, *J* = 8.04 Hz, 2H), 7.30 – 7.19 (m, 3H), 5.32 (s, 2H), 2.45 (s, 3H).

2-chlorobenzyl 4-methylbenzenesulfonate (4-5b): Yield = 76%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.82 (d, *J* = 8.40 Hz, 2H), 7.40 – 7.21 (m, 6H), 5.16 (s, 2H), 5.16 (s, 3H).

2-methylbenzyl 4-methylbenzenesulfonate (4-6b): Yield = 96%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.79 (d, *J* = 8.22 Hz, 2H), 7.32 (d, *J* = 8.04 Hz, 2H), 7.26 – 7.13 (m, 5H), 5.08 (d, *J* = 2.31 Hz, 2H), 2.45 (s, 3H), 2.25 (s, 3H).

1-methyl-3-phenylpropyl (4-methylbenzene)sulfonate (4-7b): Crude Yield = 60%.

4-methylbenzyl 4-methylbenzenesulfonate (4-8b): Yield = 99%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.79 (d, *J* = 8.22 Hz, 2H), 7.32 (d, *J* = 8.04 Hz, 2H), 7.26 – 7.13 (m, 5H), 5.08 (d, *J* = 2.31 Hz, 2H), 2.45 (s, 3H), 2.25 (s, 3H).

3,5-dimethoxybenzyl 4-methylbenzenesulfonate (4-9b): Yield = 93%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.79 (d, *J* = 8.37 Hz, 2H), 7.32 (d, *J* = 8.04 Hz, 2H), 6.39 – 6.36 (m, 3H), 4.98 (s, 2H), 3.74 (s, 6H), 2.44 (s, 3H).

3-phenylpropyl 4-methylbenzenesulfonate (4-10b): Yield = 93%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.81-7.78(d, *J* = 8.37Hz, 2H), 7.36-7.33(d, *J* = 8.01 Hz, 2H), 7.28-7.15 (m, 3H), 7.08-7.06(d, *J* = 6.96 Hz 2H), 4.06-5.01(t, *J* = 6.09 Hz, 2H), 2.67-2.62(t, *J* = 7.66 Hz, 2H), 2.5(s, 3H), 2.01-1.91(p, 2H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 145.39, 141.03, 130.51, 129.12, 129.06, 128.57, 126.80, 70.26, 22.30.

General procedure for the MW-assisted conversion of tosylates (4-1b to 4-10b) into azides (4-1, 4-2, 4-4, 4-7 and 4-11 to 4-16):

Tosylate (**4-1b** to **4-10b**) (0.4 mmol) was dissolved in DMF (2 ml) and NaN_3 (0.8 mmol) was suspended into it. The reaction mixture was pre-stirred for 30 seconds in a sealed microwave reactor and was allowed to run at 80 °C for 10 minutes under microwave irradiation, after which the reaction mixture was poured into water (10 ml) and ether (10 ml). The layers were separated and the aqueous layer was extracted with ether (2 x 10ml). The combined ether extracts were washed with brine (1 x 10ml) and dried over Na_2SO_4, filtered and concentrated under reduced pressure to obtain spectroscopically pure product (**4-1**, **4-2**, **4-4**, **4-7**, **4-11** to **4-16**, 83-98% yield) (**Fig 7.13**) which were suitable for the click reaction without any further purifications.

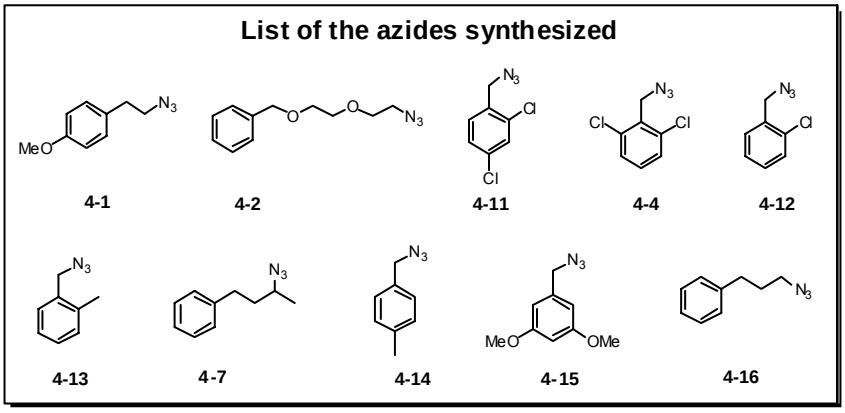


Fig. 7.13 List of azides synthesized via tosylates

1-(2-Azidoethyl)-4-methoxybenzene (4-1): Yield = 93%. 7.13-7.11(d, J = 8.55 Hz, 2H), 6.86-6.83(d, J = 8.55 Hz, 2H), 3.77(s, 3H), 3.46-3.42(t, J = 7.24 Hz, 2H), 2.84-2.79(t, J = 7.23, 2H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 159.03, 130.29, 114.60, 55.75, 53.23, 35.01.

[2-(2-Azido-ethoxy)ethoxymethyl]-benzene (4-2): Yield = 89%.

1-(azidomethyl)-2,4-dichlorobenzene (4-11): Yield = 96%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.44 – 7.42 (m, 1H), 7.35 – 7.28 (m, 2H), 4.46 (s, 2H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 134.8, 134.3, 131.9, 130.6, 129.6, 127.4, 51.6.

2-(azidomethyl)-1,3-dichlorobenzene (4-4): Yield = 98%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.37 – 7.34 (m, 2H), 7.25 – 7.19 (m, 1H), 4.67 (s, 2H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 136.3, 131.4, 130.2, 128.5, 49.1

1-(azidomethyl)-2-chlorobenzene (4-12): Yield = 93%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.46 – 7.37 (m, 2H), 7.32 – 7.24 (m, 2H), 4.48 (s, 2H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 133.8, 133.3, 130.0, 129.8, 129.7, 127.2, 52.3.

1-(azidomethyl)-2-methylbenzene (4-13): Yield = 89%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.28 – 7.22 (m, 4H), 4.35 (s, 2H), 2.37 (s, 3H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 136.8, 130.7, 130.5, 129.3, 128.6, 126.2, 53.1, 19.0.

3-azidobutyl-benzene (4-7): Crude Yield = 60%.

1-(azidomethyl)-4-methylbenzene (4-14): Yield = 93%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.25 – 7.18 (m, 4H), 4.29 (s, 2H), 2.37 (s, 3H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 129.4, 129.3, 128.5, 128.2, 54.5, 21.1.

1-(azidomethyl)-3,5-dimethoxybenzene (4-15): Yield = 96%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 6.46 – 6.38 (m, 3H), 4.27 (s, 2H), 3.80 (s, 6H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 161.1, 137.6, 106.1, 100.2, 55.4, 54.9.

(3-Azidopropyl- benzene (4-16): Yield = 91%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.40-7.35(m, 2H), 7.30-7.25(m, 3H), 3.36-3.32 (t, *J* = 6.82 Hz, 2H), 2.80-2.75(t, *J* = 7.56 Hz, 2H), 2.03-1.93(p, 2H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 141.37, 129.01, 128.95, 126.63, 51.12, 33.25, 30.94.

General procedure for the conversion of alcohols (4-a, 4-b, 4-e to 4-g, 4-q to 4-u) in to mesylates (4-1c to 4-10c)

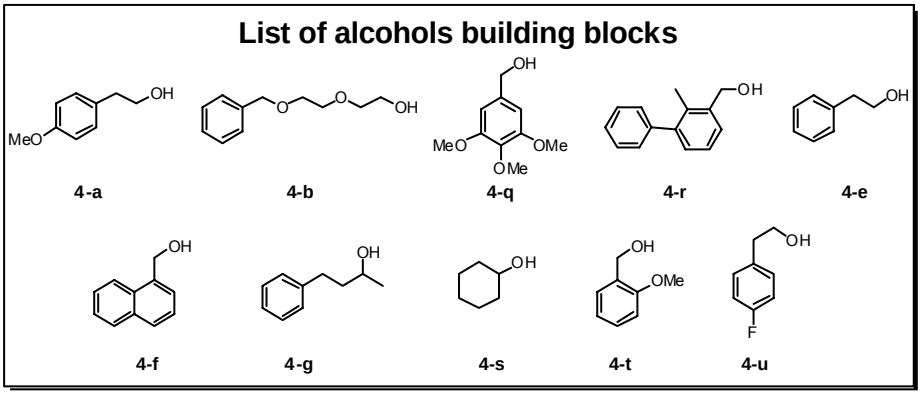


Fig. 7.14 List of alcohols used for mesylation

To a solution of the alcohol (**4-a**, **4-b**, **4-e** to **4-g**, **4-q** to **4-u**) (2 mmol) (**Fig. 7.14**) in DCM (10 ml) was added mesyl chloride (6 mmol) under ice-cold condition followed by the addition of TEA (4 mmol). The reaction mixture was stirred typically from 10 min to

one hour depending on the alcohol used. Then reaction mixture was poured into water (10 ml). The layers were separated and the aqueous layer was extracted twice with DCM (2 x 15 ml). The combined DCM extracts were washed with $NaHCO_3$ (2 x 15 ml), H_2O (1 x 15ml), brine (1 x 15 ml) and then dried over Na_2SO_4, filtered and concentrated under reduced pressure to obtain pure product. Alcohol (**4-a**, **4-b**, **4-q**, **4-e**, **4-g**, **4-s**, **4-u**) gave exclusively the respective mesylate (**4-1c** to **4-3c**, **4-5c**, **4-7c**, **4-8c**, **4-10c**) whereas alcohol (**4-r**, **4-f**, **4-t**) gave the chloride (**4-4c**, **4-6c**, **4-9c**) respectively (72-90 % yield) (**Fig. 7.15**).

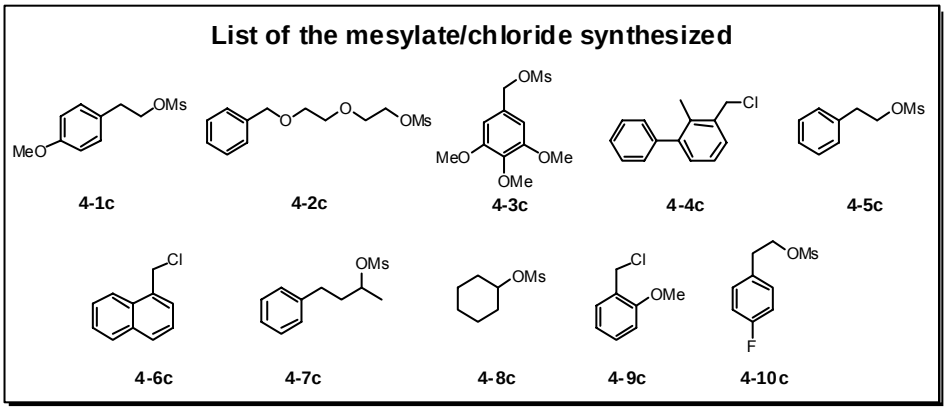


Fig. 7.15 List of mesylates/chlorides synthesized

4-methoxyphenethyl methanesulfonate (4-1c): Yield = 90%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.15 (d, *J* = 8.34 Hz, 2H), 6.86 (d, *J* = 8.70 Hz, 2H), 4.38 (t, *J* = 6.96 Hz, 2H), 3.79 (s, 3H), 3.02 (t, *J* = 6.96 Hz, 2H), 2.85 (s, 3H).

2-(2-benzylox-ethoxy-ethyl methanesulfonate (4-2c): Yield = 89%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.33-7.27(m, 5H), 4.53(s, 2H), 4.36-4.33(m, 2H), 3.71-3.78(m, 2H), 3.68-3.60(m, 6H), 2.98(s, 3H).

3,4,5-trimethoxy-benzyl methanesulfonate (4-3c): Yield = 77%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 6.62(s, 2H), 4.52(s, 2H), 3.88(s, 6H), 3.84(s, 3H).

3-(chloromethyl)-2-methylbiphenyl (4-4c): Yield = 82%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.37 – 7.20 (m, 8H), 4.58 (s, 2H), 4.68 (s, 2H), 2.31 (s, 3H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 143.3, 141.8, 136.1, 134.8, 130.7, 129.3, 129.1, 128.1, 126.9, 125.8, 45.5, 16.1.

2-phenethyl methanesulfonate (4-5c): Yield = 91%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.36-7.23 (m, 5H), 4.44-4.39(t, *J* = 6.54 Hz, 2H), 3.07-3.02(t, *J* = 6.96 Hz, 2H), 2.82(s, 3H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 136.94, 129.56, 129.27, 127.63, 70.98, 37.77, 36.14.

1-Chloromethyl-naphthalene (4-6c): Yield = 81%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.13-8.10(d, *J* = 8.37 Hz, 1H), 7.86-7.79(m, 2H), 7.59-7.46 (m, 3H), 7.41-7.35(m, 1H), 5.00(s, 2H).

1-methyl-3-(phenyl)-propyl methanesulfonate (4-7c): Yield = 72%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.22-7.18(2H), 7.04-7.00(m, 3H), 4.41(m, 1H), 3.05(s, 3H), 2.61(m, 2H), 1.79 (m, 2H), 1.38(d, 2H).

cyclohexyl methanesulfonate (4-8c): Yield = 84%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 4.73-4.64(septet, 1H), 3.00 (s, 3H), 1.97 (m, 2H), 1.79(m, 2H), 1.70-1.50(m, 3H), 1.46-1.26(m, 3H).

1-(chloromethyl)-2-methoxybenzene (4-9c): Yield = 90%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.33 – 7.25 (m, 2H), 6.94 – 6.85 (m, 2H), 4.63 (s, 2H), 3.83 (s, 3H).

2-(4-fluoro-phenyl)-ethyl methanesulfonate (4-10c): Yield = 91%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.36-7.23 (m, 5H), 4.44-4.39(t, *J* = 6.54 Hz, 2H), 3.07-3.02(t, *J* = 6.96 Hz, 2H), 2.82(s, 3H).

General procedure for the MW assisted conversion of chlorides/mesylates (4-1c to 4-10c) in to azides (4-1 to 4-2, 4-5 to 4-7 and 4-17 to 4-21)

To a solution of chloride/mesylate (1.7 mmol) in DMF (2 ml) was added NaN_3 (3.4 mmol). The reaction mixture was pre-stirred for 30 seconds in a sealed microwave reactor and was microwave irradiated for 10-15 min at 80^o (10-15 min), after which the reaction mixture was poured into water (10 ml) and ether (15 ml). The aqueous layer was extracted with ether (2 x 15 ml). The combined ether extracts were washed with $NaHCO_3$ (2 x 15 ml), water (1 x 15 ml) and brine (1 x 15 ml) and dried over Na_2SO_4, filtered and concentrated under reduced pressure to obtain a spectroscopically pure azide product (**4-1, 4-2**, **4-5** to **4-7** and **4-17** to **4-21**, 81-99 % Yield).

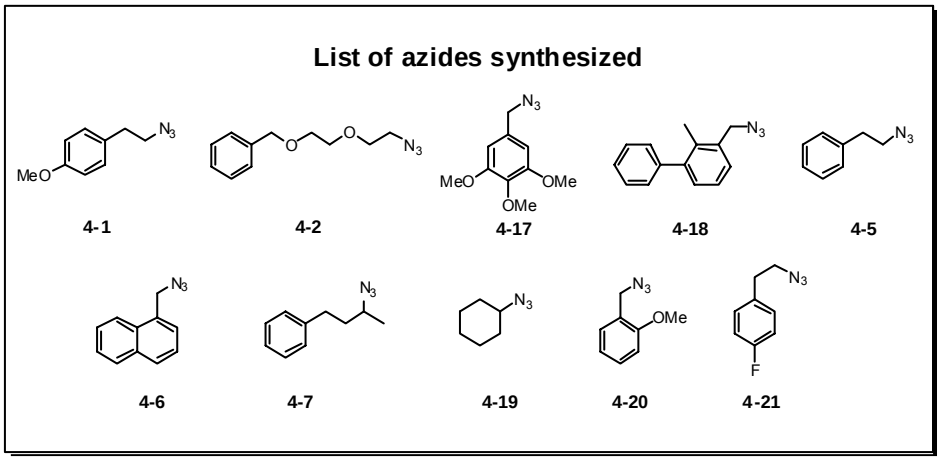


Fig. 7.16 List of azides synthesized via mesylation

1-(2-azidoethyl)-4-methoxybenzene (4-1): Yield = 99%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.13 (d, *J* = 8.73 Hz, 2H), 6.85 (d, *J* = 8.70 Hz, 2H), 3.78 (s, 3H), 3.45 (t, *J* = 7.14 Hz, 2H), 2.83 (t, *J* = 7.14 Hz, 2H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 158.5, 130.1, 129.7, 114.1, 55.2, 52.7, 34.5

[2-(2-Azido-ethoxy)-ethoxymethyl]-benzene (4-2): Yield = 87%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.35 – 7.28 (m, 5H), 4.58 (s, 2H), 3.70 – 3.64 (m, 6H), 3.40 (t, *J* = 5.09 Hz, 2H). ^{13}C-NMR: (75 MHz, $CDCl_3$) δ 138.18, 128.35, 127.68, 127.58, 73.31, 70.73, 70.03, 69.50, 50.72.

5-(azidomethyl)-1,2,3-trimethoxy-benzene (4-17): Yield = 89%. ^{1}H-NMR: (300 MHz, $CDCl_3$) δ 6.53(s, 2H), 4.28(s, 2H), 3.88(s, 6H), 3.85(s, 3H).

3-(azidomethyl)-2-methylbiphenyl (4-18): Yield = 95%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.34 – 7.16 (m, 8H), 4.38 (s, 2H), 4.68 (s, 2H), 2.21 (s, 3H); ^{13}C-NMR (75 MHz, $CDCl_3$) δ 143.2, 141.7, 134.2, 134.0, 130.3, 129.2, 128.4, 128.0, 126.9, 125.6, 53.6, 16.2.

(2-Azidoethyl)-benzene (4-5): Yield = 93%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.41-7.27(m, 5H), 3.56-3.51(t, J = 7.3 Hz, 2H), 2.96-2.91(t, *J* = 7.3 Hz, 2H), 2.96-2.92(t, *J* = 7.14Hz, 2H). ^{13}C-NMR: (75 MHz, $CDCl_3$) δ 139, 129.53, 129.28, 127.28, 52.92, 35.83.

1-(azidomethyl)-naphthalene (4-6): Yield = 81%. ^{1}H-NMR: (300 MHz, $CDCl_3$) δ 8.02-7.00(d, *J* = 8.01 Hz, 7.89-7.82(m, 2H), 7.5807.53(m, 2H), 7.44-7.42(m, 2H), 4.73(m, 2H).

3-azidobutyl-benzene (4-7): Yield = 82%. ^{1}H-NMR: (300 MHz, $CDCl_3$) δ 7.27-7.23(m, 2H), 7.18-7.14(m, 3H), 3.41-3.35(s, 1H), 2.74-2.59(m, 2H), 1.80-1.69(m, 2H), 1.2501.22(d, *J* = 6.63Hz, 3H). ^{13}C-NMR: (300 MHz, $CDCl_3$) δ 162.43, 141.23, 128.47, 128.39, 126.02, 57.14, 37.88, 32.30, 19.37.

Azidocyclohexane (4-19): Yield = 96%. ^{1}H-NMR: (300 MHz, $CDCl_3$) δ 3.36-3.30(m, 1H), 1.91-1.89(m, 2H), 1.77-1.75(m, 2H), 1.59-1.56(m, 1H), 1.43-1.21(m, 5H). ^{13}C-NMR: (300 MHz, $CDCl_3$) δ 59.97, 31.68, 25.35, 24.29

1-(azidomethyl)-2-methoxybenzene (4-20): Yield = 92%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.33 – 7.22 (m, 2H), 6.96 – 6.88 (m, 2H), 4.33 (s, 2H), 3.82 (s, 3H).

1-(2-Azido-ethyl)-4-fluoro-benzene (4-21): Yield = 94%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.41-7.27(m, 5H), 3.56, 2.96-2.91(t, J = 7.3Hz, 2H), 2.96-2.92(t, *J* = 7.14 Hz, 2H).

General procedure for the Microwave assisted click reaction:

Alkyne (0.018 mmol) was dissolved in small volume of DMSO (0.1 ml) and diluted with t-BuOH:Water (1:1) (1 ml), followed by the addition of the azide (0.020 mmol), Sodium ascorbate (0.014 mmol) and $CuSO_4.5H_2O$ (0.007 mmol). The mixture was microwave irradiated at 80^{o}C for 40 min. After which the reaction mixture was analyzed by HPLC, which indicated the complete conversion of the starting materials to yield the triazole product. Triazole products **4-T1** and **4-T3** were prepared in quantitative yields from azides **4-2** and **4-6** respectively.

Procedure for the general click reaction:

Alkyne **2-A** (0.018 mmol) was dissolved in small volume of DMSO (0.1 ml) and diluted with DCM:Water (1:1) (1 ml), followed by the addition of the azide **4-17** (0.020 mmol), Sodium ascorbate (0.0036 mmol) and $CuSO_4.5H_2O$ (0.0018 mmol). The mixture was shaken for 2 days after which it was analyzed by LC-MS, which indicated the complete conversion formation of the triazole product **4-T2.** ESI-MS: *m/z* $[M+H]^+$ = 479.149.

7.5 Activity-based fingerprinting of Enzymes

7.5.1 Experimental details of the synthesis of the probes

General procedure for synthesis of (5-23) with different amino acids:

The desired side chain-protected amino acid (1 mmol), containing either an N-α-Fmoc or N-α-Boc group, and HATU (0.456 g, 1.2 mmol) was dissolved in DMF. 2,4,6-Collidine (159 µl, 1.2 mmol) was added dropwise via a syringe. The resulting solution was agitated for 10 min before compound **5-22** (0.181 g, 1 mmol) was added. The mixture was further agitated for another 12 h, after which the DMF was removed *in vacuo* and the residue was taken into ethyl acetate (50 ml). The organic layer was washed with $NaHCO_3$ (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml) and dried over anhydrous Na_2SO_4. The organic layer was then concentrated under reduced pressure to afford a yellow oily product. Purification of this compound by flash chromatography (silica gel, ethyl acetate/hexane = 3:1) furnished **5-23a to 5-23p** typically as a pale white solid.

{4-[Boc-Ala-amino]-phenyl}-hydroxy-acetic acid methyl ester (**5-23a)**: Yield = 82%; ^{1}H-NMR (300 MHz, $CDCl_3$) 7.44 (d, *J* = 8.43 Hz, 2H), 7.26 (d, 2H), 5.11 (s, 1H), 4.37 (s, 1H), 3.71 (s, 3H), 1.43 (m, 12H); ESI-MS: *m/z* 353.1$[M + H]^+$.

{4-[Boc-Ile-amino]-phenyl}-hydroxy-acetic acid methyl ester (**5-23b)**: Yield = 87%; ESI-MS: *m/z* 417.3 $[M + Na]^+$

[4-(Boc-Phe-amino)-phenyl]-hydroxy-acetic acid methyl ester (**5-23c):** Yield = 80%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.48-7.12 (m, 9H), 5.12 (s, 1H), 4.95 (m, 1H), 3.77 (s, 3H), 3.18 (m, 2H), 1.41 (s, 9H); ESI-MS: *m/z* 429.2 $[M + H]^+$.

{4-[Fmoc-Asp(tBu)-amino]-phenyl}-hydroxy-acetic acid methyl ester (**5-23d)**: Yield = 81%; ESI-MS: *m/z* 597.4 $[M + Na]^+$.

{4-[Fmoc-Lys(Boc)amino]-phenyl}-hydroxy-acetic acid methyl ester (5-23e): Yield = 84%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.91 (s, 1H), 7.72 (d, *J* = 7.6 Hz, 2H), 7.50 (m, 4H), 7.42 (m, 6H), 5.99 (s, 1H), 5.12 (s, 1H), 4.36 (m, 2H), 4.13 (m, 2H), 3.69 (m, 3H), 3.05 (m, 2H), 1.90 (m, 1H), 1.70 (m, 1H), 1.41 (m, 11H), 1.26 (m, 2H); ESI-MS: *m/z* 654.3 $[M + Na]^+$.

{4-[Fmoc-Thr(tBu)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-23f): Yield = 84%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.76 (d, 2H), 7.58 (m, 4H), 7.33 (m, 6H), 6.08 (m, 1H), 5.15 (s, 1H), 4.42 (m, 2H), 4.24 (m, 1H), 3.77 (m, 1H), 3.75 (s, 3H), 1.35 (s, 9H), 1.26 (m, 3H); ESI-MS: *m/z* 583.3 $[M + Na]^+$.

{4-[Fmoc-Tyr(tBu)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-23g): Yield = 82%; ^{1}H-NMR (300 MHz, $CDCl_3$) 7.75 (d, *J* = 7.62 Hz, 2H), 7.38 (m, 4H), 7.27 (m, 6H), 7.19 (m, 2H), 6.89 (d, *J* = 8.04 Hz, 2H), 5.12 (s, 1H), 4.92 (s, 1H), 4.41 (m, 2H), 4.19 (m, 1H), 3.73 (s, 3H), 3.19 (m, 2H), 1.3(s, 9H); ESI-MS: *m/z* 623.4$[M + H]^+$.

{4-[Fmoc-Glu(OBut-*t*)amino]-phenyl}-hydroxy-acetic acid methyl ester (5-23h): Yield = 85%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.86 (s, 1H), 7.74 (d, *J* = 7.6 Hz, 2H), 7.54 (m, 4H), 7.30 (m, 6H), 6.08 (s, 1H), 5.14 (s, 1H), 4.38 (m, 3H), 4.18 (m, 1H), 3.71 (m, 3H), 2.41 (m, 2H), 2.04 (m, 2H), 1.44 (m, 9H); ESI-MS: *m/z* 611.1 $[M + Na]^+$.

{4-[Fmoc-Asn(Trt)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-23i):

Yield = 84%. ESI-MS: *m/z* 761.6 $[M + H]^+$

{4-[Fmoc-Gln(Trt)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-23j): Yield = 84%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.73 (d, *J* = 7.23 Hz, 2H), 7.47 (m, 4H), 7.26 (m, 21H), 5.84 (s, 1H), 4.36 (m, 2H), 4.16 (m, 2H), 3.69 (m, 3H), 2.41 (m, 2H), 2.07 (m, 2H); ESI-MS: *m/z* 774.2 $[M + H]^+$.

{4-[Fmoc-Trp(Boc)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-23k): Yield = 80%; ESI-MS: *m/z* 712.8 $[M + Na]^+$.

{4-[Fmoc-Val-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-23l): Yield = 86%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.71 (d, 2H), 7.52 (m, 4H), 7.26 (m, 6H), 5.12 (s, 1H), 4.38 (m, 2H), 4.21 (m, 1H), 4.11 (m, 1H), 3.68 (s, 3H), 2.13 (m, 1H), 1.02 (d, 6H); ESI-MS: *m/z* 525.7$[M + Na]^+$.

{4-[Fmoc-His(Trt)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-23m): Yield = 86%. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.75 (d, 2H), 7.51 (m, 4H), 7.29 (m, 21H), 7.02 (m, 1H), 6.68 (m, 1H), 5.16 (s, 1H), 4.65 (m, 1H), 4.38 (m, 2H), 4.12 (m, 1H), 3.71 (m, 3H), 3.16(m, 2H); ESI-MS: *m/z* 782.9 $[M + H]^+$.

{4-[Fmoc-Ser(tBu)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-23n): Yield = 83%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.75 (d, 2H), 7.60 (m, 4H), 7.38 (m, 6H), 5.16 (s,

1H), 4.44 (m, 2H), 4.24 (m, 1H), 4.10 (m, 1H), 3.71 (s, 3H), 3.48 (m, 2H), 1.23 (s, 9H); ESI-MS: *m/z* 547.3$[M + H]^+$.

{4-[Fmoc-Leu-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-23o): Yield = 88%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.73 (d, 2H), 7.55 (m, 4H), 7.28 (m, 6H), 5.13 (s, 1H), 4.40 (m, 2H), 4.23 (m, 1H), 4.14 (m, 1H), 3.69 (s, 3H), 1.71 (m, 3H), 0.91 (m, 6H); ESI-MS: *m/z* 539.1$[M + Na]^+$.

{4-[Fmoc-Cys(Trt)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-23p): Yield = 82%; ESI-MS: *m/z* 717.6$[M + Na]^+$.

General procedure for synthesis of (5-24) with different amino acids:

Depending upon the α-amino protecting group, 20% Piperidine in DMF or 50% TFA in DMF (5 ml) was added to compound **5-23** (0.5 mmol) and the resulting solution was stirred for 30 min. Following removal of the solvent *in vacuo*, the residue was dissolved in anhydrous DCM and a mixture of acetic anhydride (0.153 g, 1.5 mmol) and DIEA (0.194 g, 1.5 mmol) was added. After 5 min, the DCM was removed and the residue was purified by flash chromatography (silica gel, ethyl acetate/hexane = 3:1) to furnish **5-24a to 5-24p** typically as a dull white solid.

{4-[Acetyl–Ala-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24a):

Yield = 56%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.44 (d, 2H), 7.26 (d, 2H), 5.11 (s, 1H), 4.67 (s, 1H), 3.71 (s, 3H), 2.03 (s, 3H), 1.42 (m, 3H); ESI-MS: *m/z* 295.4 $[M + H]^+$.

{4-[Acetyl–Ile-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24b): Yield = 54%; ESI-MS: *m/z* 359.3 $[M + H]^+$.

[4-(Acetyl-Phe-amino)-phenyl]-hydroxy-acetic acid methyl ester (5-24c): Yield = 45%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.48-7.12 (m, 9H), 5.12 (s, 1H), 4.95 (m, 1H), 3.77 (s, 3H), 3.18 (m, 2H), 1.99 (s, 3H); ESI-MS: *m/z* 370.9 $[M+H]^+$.

{4-[Acetyl-Asp(O-But)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24d): Yield = 53%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.51 (d, 2H), 7.32 (d, 2H), 5.11 (s, 1H), 4.89 (m, 1H), 3.69 (s, 3H), 2.85 (m, 2H), 2.00 (s, 3H), 1.49 (s, 9H); ESI-MS: *m/z* 395.1 $[M + H]^+$.

{4-[Acetyl-Lys(Boc)amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24e): Yield = 53%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.45 (d, J = 8.4 Hz, 2H), 7.29 (d, J = 8.43 Hz, 2H), 5.12 (s, 1H), 4.53 (m,1H), 3.67 (s, 3H), 3.00 (m, 2H), 2.03 (s, 3H), 1.90 (m, 2H), 1.41 (m, 11H), 1.26 (m, 2H); ESI-MS: *m/z* 452.1 $[M+H]^+$.

{4-[Acetyl-Thr(t-But)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24f): Yield = 56%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.48 (d, 2H), 7.31 (d, 2H), 5.13 (s, 1H), 4.46

(m, 1H), 3.98 (m, 1H), 3.71 (s, 3H), 2.03 (s, 3H), 1.39 (s, 9H), 1.28 (s, 3H); ESI-MS: *m/z* 380.1 $[M + H]^+$.

{4-[Acetyl–Tyr(Bu-t)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24g): Yield = 50%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.36 (d, J = 8.40 Hz, 2H), 7.22 (d, J = 8.01 Hz, 2H), 7.11 (d, *J* = 8.01 Hz, 2H), 6.85 (d, *J* = 8.01 Hz, 2H), 5.11 (s, 1H), 4.99 (m, 1H), 3.68 (s, 3H), 3.17 (m, 2H), 1.9 (s, 3H), 1.24 (s, 9H); ESI-MS: *m/z* 465.4 $[M + Na]^+$.

{4-[Acetyl-Glu(OBut-*t*)amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24h): Yield = 50%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.49 (d, *J* = 8.43 Hz, 2H), 7.37 (d, *J* = 8.43 Hz, 2H), 5.12 (s, 1H), 4.56 (m, 1H), 3.67 (s, 3H), 2.31 (m, 4H), 2.04 (s, 3H), 1.44 (s, 9H); ESI-MS: *m/z* 409.2 $[M+H]^+$.

{4-[Acetyl–Asn(Trt)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24i): Yield = 53%; ESI-MS: *m/z* 580.5 $[M + H]^+$.

{4-[Acetyl-Gln(Trt)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24j): Yield = 54%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.73 (d, *J* = 78.23 Hz, 2H), 7.49 (m, 2H), 7.27 (m, 15H), 5.83 (s, 1H), 4.41 (m, 1H), 3.67 (m, 3H), 2.02 (s, 3H), 2.43 (m, 2H), 2.12 (m, 2H); ESI-MS: *m/z* 594.9 $[M + H]^+$.

{4-[Acetyl–Trp(Boc)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24k): Yield = 49%.

{4-[Acetyl–Val-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24l): Yield = 55%; ESI-MS: *m/z* 323.1 $[M + H]^+$.

{4-[Acetyl–His(Trt)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24m): Yield = 53%; ESI-MS: *m/z* 601.7 $[M +1]^+$.

{4-[Acetyl-Ser(tBu)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24n): Yield = 53%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.37 (d, *J* = 8.43 Hz, 2H), 7.29 (s, *J* = 8.43 Hz, 2H), 5.16 (s, 1H), 4.61 (m, 1H), 3.70 (s, 3H), 3.49 (m, 2H), 2.06 (s, 3H), 1.24 (S, 9H); ESI-MS: *m/z* 367.1$[M + H]^+$.

{4-[Acetyl-Leu-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24o): Yield = 53%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.48 (d, 2H), 7.35 (d, 2H), 5.13 (s, 1H), 4.53 (m, 1H), 3.68 (s, 3H), 2.0 (s, 3H), 1.70 (m, 3H), 0.90 (m, 6H); ESI-MS: *m/z* 359.1$[M + Na]^+$.

{4-[Acetyl–Cys(Trt)-amino]-phenyl}-hydroxy-acetic acid methyl ester (5-24p): Yield = 47%.

General procedure for synthesis of compound 5-27:

To a solution of compound **5-24** (0.1 mmol) in methanol (9 ml), 0.1 M LiOH (3 ml) was added and the mixture was stirred at room temperature. After 3 h, the resulting mixture was concentrated under reduced pressure. Following addition of 1 M HCl solution (10 ml), the product was extracted with 20 ml of ethyl acetate or dichloromethane and the organic layer was washed with 1 M HCl (2 x 15 ml), water (2 x

15 ml), and brine (2 x 15 ml), and dried over anhydrous Na_2SO_4 followed by concentration *in vacuo* to afford **5-25** typically as a dull solid, which was used for subsequent steps without further purification.

A Fmoc-mono-protected diamine linker, **5-28** (44 mg, 0.12 mmol) as shown below and synthesized as previously described,[44] was added to a flask containing compound **5-25** dissolved in DMF (15 ml). Subsequently, HOBt (18 mg, 0.12 mmol) and EDC (23 mg, 0.12 mmol) were added and the reaction was allowed to proceed at 0 °C for 30 min with stirring. At the end of this period, the reaction mixture was warmed to room temperature and further stirred for an additional 8 h, after which aqueous NH_4Cl was added and the product was extracted with ethyl acetate. The combined organic extract was washed with 1 M HCl (2 x 50 ml), water (2 x 50 ml) and brine (2 x 50 ml), followed by drying over anhydrous Na_2SO_4 and concentration under reduced pressure to furnish the crude product **5-26**. Without further purification, this compound was dissolved in CH_2Cl_2 (12 ml) at 0 °C and (diethylamino)sulfur trifluoride (0.033 ml, 0.3 mmol) was added dropwise. The reaction temperature was then increased to 25 °C. After stirring for 30 min, the solvent was removed and the residue was taken into ethyl acetate (50 ml). The organic layer was washed with $NaHCO_3$ (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml) followed by drying over anhydrous Na_2SO_4 and concentration under reduced pressure to afford the crude product, which was further purified by flash chromatography (silica gel, hexane/ethyl acetate = 1:3) to afford **5-27a to 5-27p.**

Compound 5-27a (Ala): Yield = 63%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.71 (d, 2H), 7.56 (m, 4H), 7.29 (m, 6H), 5.68 (d, J = 48.6 Hz, 1H), 5.37 (m, 1H), 4.70 (m, 2H), 4.40 (m,

1H), 3.58 (s, 10 H), 3.36 (m, 2H), 2.01 (s, 3H), 1.43 (m, 3H); ESI-MS: *m/z* 657.2 [M + Na]$^+$.

Compound 5-27b (Ile): Yield = 63%; ESI-MS: *m/z* 676.8 [M + H]$^+$.

Compound 5-27c (Phe): Yield = 53%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.75 (d, *J* = 7.2 Hz, 2H), 7.68 (m, 4H), 7.36 (m, 6H), 5.70 (d, *J* = 48.6 Hz, 1H), 5.35 (m, 1H), 4.54 (m, 1H), 4.39 (m, 2H), 4.20 (m, 2H), 3.53 (m, 10H), 3.38 (m, 2H), 2.46 (m, 2H), 2.10 (m, 2H), 2.01 (s, 3H), 1.45 (m, 9H); ESI-MS: *m/z* 733.3 [M + Na]$^+$.

Compound 5-27d (Asp): Yield = 63%; ESI-MS: *m/z* 758.1 [M + Na]$^+$.

Compound 5-27e (Lys): Yield = 66%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.75 (d, *J* = 7.2 Hz, 2H), 7.58 (m, 4H), 7.36 (m, 6H), 5.70 (d, *J* = 48.5 Hz, 1H), 5.33 (m, 1H), 4.59 (m, 2H), 4.40 (m, 1H), 3.61 (m, 10H), 3.37 (m, 2H), 3.09 (m, 2H), 2.04 (s, 3H), 1.90 (m, 2H), 1.43 (m, 13H); ESI-MS: *m/z* 814.4 [M + Na]$^+$.

Compound 5-27f (Thr): Yield = 64%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.73 (d, 2H), 7.39 (m, 10H), 5.69 (d, *J* = 48.4 Hz, 1H), 5.38 (m, 1H), 4.49 (m, 2H), 4.21 (m, 1H), 3.98 (m, 1H), 3.61 (m, 10H), 3.34 (m, 2H), 2.02 (s, 3H), 1.39 (s, 9H), 1.29 (m, 3H); ESI-MS: *m/z* 744.4 [M + Na]$^+$.

Compound 5-27g (Tyr): Yield = 63%; ESI-MS: *m/z* 806.1[M + Na]$^+$.

Compound 5-27h (Glu): Yield = 60%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.75 (d, *J* = 7.2 Hz, 2H), 7.58 (d, *J* = 7.7 Hz, 2H), 7.31 (m, 12H), 7.06 (m, 1H), 5.70 (d, *J* = 48.5 Hz, 1H), 5.43 (m, 1H), 4.41 (m, 2H), 4.20 (m, 1H), 3.63 (m, 10H), 3.37 (m, 2H), 3.07 (m, 2H), 2.04 (s, 9H); ESI-MS: *m/z* 771.3 $[M + Na]^+$.

Compound 5-27i (Asn): Yield = 64%; ESI-MS: *m/z* 943.8 $[M + H]^+$.

Compound 5-27j (Gln): Yield = 67%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.72 (d, 2H), 7.58 (m, 4H), 7.27 (m, 21H), 5.72 (d, *J* = 48.5 Hz, 1H), 5.37 (m, 1H), 4.69 (m, 2H), 4.38 (m, 1H), 3.58 (s, 10H), 3.34 (m, 2H), 2.00 (s, 3H), 2.44 (m, 2H), 2.1 (m, 2H); ESI-MS: *m/z* 935.1 $[M + H]^+$.

Compound 5-27k (Trp): Yield = 63%; ^{1}H-NMR (300 MHz, $CDCl_3$) 7.75 (d, 2H), 7.57 (m, 4H), 7.39 (m, 6H), 7.20 (m, 2H), 7.01 (d, 2H), 6.90 (m, 1H), 5.71 (d, *J* = 48.6 Hz, 1H), 5.39 (m, 1H), 4.59 (m, 2H), 4.39 (m, 1H), 3.62 (m, 10H), 3.38 (m, 2H), 3.18 (m, 2H), 2.02 (s, 3H), 1.25 (s, 9H); ESI-MS: *m/z* 872.3 $[M + Na]^+$.

Compound 5-27l (Val): Yield = 59%; ESI-MS: *m/z* 663.4 $[M + H]^+$.

Compound 5-27m (His): Yield = 60%; ESI-MS: *m/z* 944.1 $[M + H]^+$.

Compound 5-27n (Ser): Yield = 61%; ESI-MS: *m/z* 706.9$[M + H]^+$.

Compound 5-27o (Leu): Yield = 64%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.70 (d, 2H), 7.56 (m, 4H), 7.35 (m, 6H), 5.79 (d, *J* = 48.5 Hz, 1H), 5.34 (m, 1H), 4.41 (m, 2H), 4.21 (m, 1H), 3.60 (m, 10H), 3.38 (m, 2H), 2.00 (s, 3H), 1.71 (m, 3H), 0.91 (m, 6H); ESI-MS: *m/z* 699.8 $[M + Na]^+$.

Compound 5-27p (Cys): Yield = 64%; ESI-MS: *m/z* 932.3 $[M + Na]^+$.

General procedure for synthesis of the Probes (5-1 to 5-16):

To compound **5-27** (0.05 mmol), 20% piperidine in DMF (3 ml) was added and the resulting solution was stirred for 30 min. After which the solvent was removed under reduced pressure to yield a residue, which upon taken into DMF (3 ml), was added Cy3-NHS (32.4 mg, 0.06 mmol) and DIEA (10 μl, 0.06 mmol). The reaction mixture was stirred at room temperature for 18 h followed by concentration under reduced pressure to give the crude product as syrup. Following de-protection by treatment with 50% TFA/DCM for 30 min, the resulting product was further purified by reverse-phase HPLC (CH_3CN/H_2O/0.01% TFA; 15/85 → 60/40 gradient in 60 min) to give the desired product (**Probes 5-1** to **5-16**). The products were analyzed by NMR and MS analysis. The results of the MS analysis are shown in **Table 7.2.**

Probe 5-1 (Ala): Yield =75%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.36 (m, 1H), 7.74 (m, 2H), 7.37 (m, 6H), 7.25 (m, 2H), 7.13 (m, 2H), 6.48 (m, 2H), 5.74-5.58 (d, *J* = 48.6 Hz, 1H), 4.62 (m, 1H), 3.98 (m, 2H), 3.56 (m, 12H), 3.40 (m, 2H), 2.01 (m, 3H), 1.90 (m, 19 H); ESI-MS: *m/z* 838.5 $[M - I]^+$.

Probe 5-2 (Ile): Yield = 70 %; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.37 (m, 1H), 7.75 (m, 2H), 7.37 (m, 6H), 7.25 (m, 2H), 7.12 (m, 2H), 6.44 (m, 2H), 5.74-5.58 (d, *J* = 48.5 Hz, 1H), 4.57 (m, 1H), 4.08 (m, 2H), 3.60 (m, 12H), 3.41 (m, 2H), 2.50 (m, 1H), 2.03 (m, 3H), 1.71 (m, 18), 0.92 (m, 6H); ESI-MS: *m/z* 880.5 $[M - I]^+$.

Probe 5-4 (Asp): Yield = 71%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.36 (m, 1H), 7.55 (d, 2H), 7.26 (m, 6H), 7.17 (m, 2H), 7.14 (m, 2H), 6.39 (m, 2H), 5.73-5.57 (d, *J* = 48.5 Hz, 1H), 4.94 (m, 1H), 3.59 (m, 2H), 3.53 (m, 12H), 3.31 (m, 2H), 2.84 (m, 2H), 2.17 (s, 3H), 1.7 (m, 16H); ESI-MS: *m/z* 882.4 $[M - I]^+$.

Probe 5-6 (Thr): Yield = 72%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.37 (m, 1H), 7.63 (m, 2H), 7.37 (m, 6H), 7.26 (m, 2H), 7.14 (m, 2H), 6.41 (m, 2H), 5.76-5.59 (d, *J* = 48.4 Hz, 1H), 4.94 (m, 1H), 4.71 (m, 1H), 3.90 (m, 2H), 3.58 (m, 12H), 3.43 (m, 2H), 2.16 (s, 3H), 1.69 (m, 16H), 1.28 (m, 3H); ESI-MS: *m/z* 868.5 $[M - I]^+$.

Probe 5-7 (Tyr): Yield = 72%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.35 (m, 1H), 7.58 (m, 2H), 7.31 (m, 2H) 7.25 (m, 6H), 7.12 (m, 2H), 6.80 (m, 2H), 6.51 (m, 2H), 6.45 (m, 2H), 5.58-5.74 (d, *J* = 48.5 Hz, 1H), 4.67 (m, 1H), 3.98 (m, 2H), 3.55 (m, 12H), 3.41 (m, 2H), 3.21 (m, 2H), 2.21 (s, 3H), 1.97 (m, 16H); ESI-MS: *m/z* 930.5 $[M - I]^+$.

Probe 5-9 (Asn): Yield = 71%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.37 (m, 1H), 7.67 (m, 2H), 7.37 (m, 8H), 7.94 (m, 2H), 6.49 (m, 2H), 5.58-5.74 (d, *J* = 48.4 Hz, 1H), 4.90 (m,

1H), 4.00 (m, 2H), 3.60 (m, 12H), 3.32 (m, 2H), 2.49 (m, 2H), 2.10 (s, 3H), 1.71 (m, 16H); ESI-MS: *m/z* 880.5 $[M - I]^+$.

Probe 5-10 (Gln): Yield = 72%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.36 (m, 1H), 7.72 (m, 2H), 7.37 (m, 8H), 7.13 (m, 2H), 6.47 (m, 2H), 5.75 (d, *J* = 48.5 Hz, 1H), 4.62 (m, 1H), 3.98 (m, 2H), 3.59 (m, 12 H), 3.34 (m, 2H), 2.29 (m, 2H), 2.18 (m, 2H), 2.04 (s, 3H), 1.74 (m, 16H); ESI-MS: *m/z* 895.4 $[M - I]^+$.

Probe 5-11 (Trp): Yield = 72%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.36 (m, 1H), 7.65 (m, 2H), 7.47 (m, 10H), 7.34 (m, 2H), 7.14 (m, 2H), 6.98 (m, 1H), 6.54 (m, 2H), 5.73-5.57 (d, *J* = 48.6 Hz, 1H), 4.86 (m, 1H), 4.10 (m, 2H), 3.52 (m, 12H), 3.30 (m, 2H), 3.01 (m, 2H), 1.98 (s, 3H), 1.70 (m, 16H); ESI-MS: *m/z* 953.4 $[M - I]^+$.

Probe 5-12 (Val): Yield = 74%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.33-8.41 (m, 1H), 7.68-7.71 (m, 2H), 7.40-7.33 (m, 8H), 7.27 (m, 2 H), 6.41 (m, 2H), 5.59-5.75 (d, *J* = 48.5 Hz, 1H), 4.39-4.51 (m, 1H), 3.99 (m, 2H), 3.64 (m, 12H), 3.31 (m, 2H), 2.29 (m, 1H), 2.04 (s, 3H), 1.74 (m, 16H), 0.97 (m, 6H); ESI-MS: *m/z* 866.4 $[M - I]^+$.

Probe 5-13 (His): Yield = 72%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.39 (m, 1H), 7.71 (m, 3H), 7.38 (m, 6H), 7.19 (m, 2H), 7.13 (m, 3H), 6.36 (m, 2H), 5.76-5.60 (d, *J* = 48.6 Hz, 1H), 4.94 (m, 1H), 4.00 (m, 2H), 3.58 (m, 12H), 3.15 (m, 2H), 2.95 (m, 2H), 2.09 (s, 3H), 1.72 (m, 16H); ESI-MS: *m/z* 904.5 $[M - I]^+$.

Probe 5-14 (Ser): Yield = 69%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.37 (m, 1H), 7.66 (m, 2H), 7.36 (m, 6H), 7.22 (m, 2H), 7.12 (m, 2H), 6.40 (m, 2H), 5.59-5.75 (d, *J* = 48.5 Hz, 1H), 4.63 (m, 1H), 4.12 (m, 2H), 3.97 (m, 2H), 3.59 (m, 12H), 3.37 (m, 2H), 2.05 (s, 3H), 1.71 (m, 16H); ESI-MS: *m/z* 854.5 $[M - I]^+$.

Probe 5-15 (Leu): Yield = 73%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.37 (m, 1H), 7.75 (m, 2H), 7.38 (m, 6H), 7.24 (m, 2H), 7.12 (m, 2H), 6.44 (m, 2H), 5.58-5.74 (d, *J* = 48.5 Hz, 1H), 4.58 (m, 1H), 4.10 (m, 2H), 3.61 (m, 12H), 3.40 (m, 2H), 2.04 (m, 3H), 1.71 (m, 19H), 0.91 (m, 6H); ESI-MS: *m/z* 880.5 $[M - I]^+$.

Probe 5-16 (Cys): Yield = 71%; ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.37 (m, 1H), 7.70 (d, 2H), 7.38 (m, 6H), 7.25 (m, 2H), 7.17 (m, 2H), 6.41 (m, 2H), 5.75-5.59 (d, *J* = 48.5 Hz, 1H), 4.76 (m, 1H), 3.99 (m, 2H), 3.53 (m, 12H), 3.39 (m, 2H), 3.21 (m, 2H), 2.04 (s, 3H), 1.72 (m, 16H); ESI-MS: *m/z* 870.5 $[M - I]^+$.

Probe	$[M-I]^+$	Actual Mass	probe	$[M-I]^+$	Actual
Ala	838.50	837.47	Asn	880.50	880.48
Ile	880.50	879.52	Gln	895.40	894.49
Phe	914.40	913.30	Trp	953.40	952.51
Asp	882.40	881.46	Val	866.30	866.40
Lys	894.60	894.53	His	904.50	903.60
Thr	868.50	867.48	Ser	854.50	853.47
Tyr	930.50	929.50	Leu	880.50	879.52
Glu	895.60	895.48	Cys	870.50	869.44

Table 7.2 Summary of ESI-MS results of the probes **5-1** to **5-16**.

7.5.2 Gel-Based Labeling Experiments with 16 Probes:

Stock solutions of proteins were prepared as ~ 2 mg/ml solutions in distilled water, desalted with a NAP5 column (Amersham, USA) and stored at -20 ^{o}C until use. Since the different proteins, as obtained from commercial sources, contain different amount of salts, the final absolute weights of the protein in the estimated 2 mg/ml stock solution likely differ to some degree (as determined by coomassie staining of the proteins upon separation on a SDS-PAGE gel). Furthermore, the relative catalytic activity of each commercial protein also differs as well. Both factors render the quantitation of our labeling reactions across different proteins less relevant. Consequently, the relative fluorescence labeling in each protein is more meaningful when the labeling of the same protein against different probes is compared. The labeling of different proteins by the

same or different probes may only be used as a qualitative measure. The stock solutions of the 16 probes were prepared as 200 μM solutions in methanol and stored at -20 ^{o}C. Fluorescence imaging was carried out using the Typhoon 9200 scanner (Amersham, USA).

(Enzyme) Concentration-dependent labeling

Probe 5-5 (2 μM) was incubated with different concentration of trypsin using the procedure as described above. The sample was then resolved on a 12% gel and visualized by fluorescence scanning. It was found (**Fig. 7.17**) that as low as 1 ng of the enzyme could be detected with μM of the probe **5-5**.

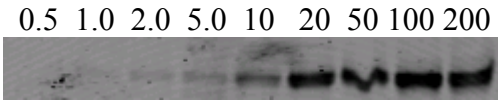


Fig. 7.17 Increasing concentrations of trypsin labeled

(Probe) Concentration-dependent labeling

Reactions containing 0.02 mg/ml of trypsin were treated with varied concentrations of **Probe 5-5** (0.05-2 μM) using the above descried procedure. The resulting samples were resolved on a 12% gel and visualized by fluorescence scanning. As shown in **Fig 7.18**, as low as 0.1 μM of the probe is needed for protein labeling to be detected.

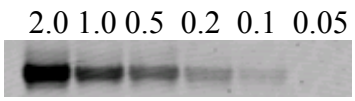


Fig. 7.18 Trypsin (0.02 mg/ml) labeled with increasing concentrations of **Probe 5-5**

Fingerprinting experiments

Thirteen different proteases were individually labeled with Probes **5-1** to **5-16** as described above. The resulting gel-based profiles (**Fig. 7.19**) were used to generate fingerprints of the enzymes as shown. With trypsin as a representative, a unique activity-based fingerprinting profile of the enzyme against all probes was thus generated. Based on the relative intensity of the labeled enzymes by each probe, a dendrogram were constructed using the programs Tree View™ and Cluster™ developed by Eisen and coworkers. The thirteen proteases were hierarchically clustered based on the similarity of their fingerprint profiles, and shown by the tree structure on the left of the dendrogram (**Fig. 5.4**). The degree of similarity of protease was indicated as a function of height of the lines connecting profiles.

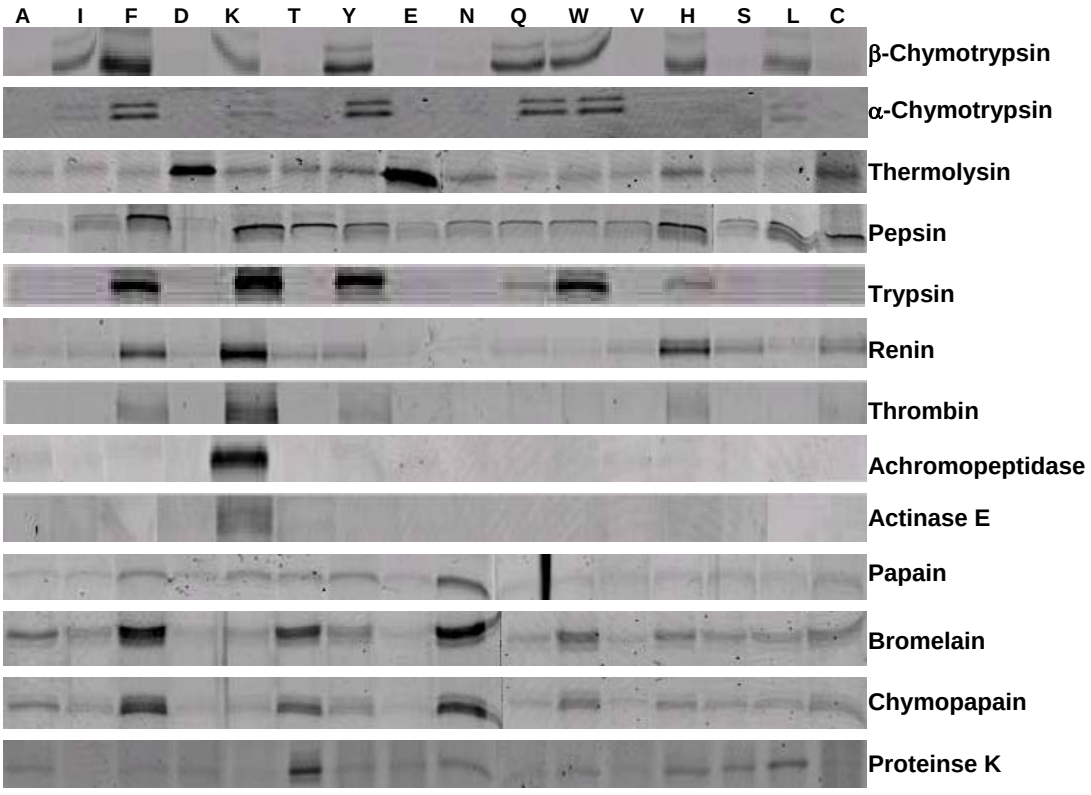


Fig. 7.19 SDS-PAGE of 13 different proteases labeled with **Probes 5-1** to **5-16**. Each lane in the gel represents labeling by a different probe (from left to right; shown with three-letter amino acid code): Ala, Ile, Phe, Asp, Lys, Thr, Tyr, Glu, Asn, Gln, Trp, Val, His, Ser, Leu, Cys.

7.5.3 Chemical Synthesis of the PTP probes

***p*-Hydroxymandelic acid (5-19):** A solution of phenol (11.3 g , 0.12 mol), glyoxalic acid (7.4 g, 0.1 mol), and NaOH (8 g) was dissolved in 80 ml of water and heated to 40˚C for 8 hrs. Then the reaction solution was brought to a pH of 6 using 1N HCl then the unconverted phenol was extracted with benzene (3 x 100 ml). The aqueous solution was acidified to a pH of 1.5 using 1N HCl and extracted with ethyl acetate (10 x 50 ml) and washed with brine (1 x 100 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated to 30% of the weight and left to crystallize over night at 0 ˚C to afford pure **5-19** (12.1 g, 60%) in the form of solid prisms. ^{1}H-NMR (300 MHz, DMSO) δ 7.25 (d, *J* = 3.0 Hz, 2H), 6.84 (d, *J* = 3.0 Hz, 2H). MS (ESI): m/z 191.4 $[(M+23)]^+$.

[2-(2-{2-[2-hydroxy-2-(4-hydroxyphenyl)-acetylamino]-ethoxy}-ethoxy-ethyl] carbamic acid tert-butyl ester (5-20): To a solution of amine (10 g, 0.059mol) in dry DMF were added EDC (13.6g, 0.071 mol), HOBt (9.94 g, 0.065 mol) and compound **5-19** (14.65 g, 0.059mol) under ice cold condition. The resulting solution was further stirred for 24 hrs at RT. Upon removal of the solvent *in vacuo*, the residue was taken in to ethyl acetate and washed with water (2 x 80 ml) and brine (2 x 80 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to afford a crude mixture, which was further purified by flash chromatography (silica gel, ethyl

acetate 100%) to afford **5-20** (18.78 g, 80 %) as a dull brown highly viscous oil. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.09(d, *J* = 3, 2H), 6.63(d, *J* = 2, 2H), 4.89(s, 1H), 3.54-3.39 (m, 10 H), 3.21-3.19(t, 2H), 1.38 (s, 9H). MS (ESI): m/z 421.3 $[(M+23)]^+$.

Compound 5-21: To a solution of **5-20** (7 g, 17.6 mmol) in dry DCM was added TEA (4.8 ml, 44 mmol) and diethyl chlorophosphate (2.4 ml, 17.6 mmol) at 0 ^oC. The resulting solution was further stirred for 24 hrs at RT. The reaction was then quenched with water and washed with saturated $NaHCO_3$ (2 x 60 ml), water (2 x 60 ml) and brine (2 x 60 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure and further purified by flash chromatography (silica gel, dichloromethane/ethanol = 8:2) to afford **5-21** (7.47 g, 80 %) as a dull brown viscous liquid. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.26(d, *J* = 3 Hz, 2H), 6.98 (d, *J* = 3 Hz, 2H), 4.85(s,1H), 4.07-3.98 (p, 4H), 3.42 (s, 4H), 3.38-3.34 (t, 4H), 3.31-3.27 (t, 2H), 3.14-3.09 (m,2H), 1.29 (3, 9H), 1.21-1.16 (t, 6H). MS (ESI): m/z 557.2 $[(M+23)]^+$.

Compound 5-22: Compound **5-21** (5 g, 9.36 mmol) was dissolved in dry DCM and DAST (3.67 ml, 28 mmol) was slowly added under ice cold condition. The mixture was allowed to stir for further 1 hour at RT. Then the reaction mixture was washed with cold water and dried with anhydrous Na_2SO_4. The organic layer was concentrated under reduced pressure and further purified by flash chromatography to afford **5-22** (4.01 g, 78%) as a dark brown oily liquid. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.39 (d, *J* = 2.8 Hz, 2H), 7.19 (d, *J* = 3 Hz, 2H), 5.78 (s, 1H), 5.61 (s, 1H), 4.21-4.10 (p, 4H), 3.56-3.47 (m, 8H), 3.26-3.24 (m, 2H), 1.38 (s, 9H), 1.27-1.31 (t, 6H) MS (ESI): m/z 559.2 $[(M+23)]^+$.

Compound 5-23: Compound **5-22** (4 g, 7.46 mmol) was added to a 40 ml solution of 1:1 DCM:TFA mixture and stirred for 30 minutes after which it was evaporated under reduced pressure to afford the amine **5-23** (2.93 g, 90 %). It was used with out further purification. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.45(d, J = 4.29 Hz, 2H), 7.19(d, *J* = 2.68 Hz, 2H) 5.87(s, 1H), 5.71(s, 1H), 4.30-4.2(p, 4H), 3.62-3.59(m, 8H), 3.53-3.45 (m, 2H), 1.36-1.31 (t, 6H). MS (ESI): m/z 437.2 $[(M+1)]^+$.

Compound 5-24: To a solution of **5-23** (2.93 g, 6.72 mmol) in dry DMF was added EDC (1.55g, 8.06 mmol), HOBt (1.13g, 7.39 mmol) and Fmoc-Lys(Boc)-OH (3.15 g, 6.72 mmol) under ice cold condition and stirred for further 12 hrs at RT. The solvent was evaporated at reduced pressure and the residue was taken in to DCM and washed with $NaHCO_3$ (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure and further purified by flash column (silica gel, dichloromethane/ethanol = 9.5:0.5) to afford **5-24** (4.52g, 76%) as a dull brownish white solid. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.74(d, *J* = 2.4 Hz, 2H), 7.55(d, *J* = 2.27 Hz, 2H), 7.41-7.37 (t, 2H), 7.21-7.28 (t, 2H), 7.19-7.21 (m, 4H), 5.83 (s, 1H), 5.67 (s, 1H), 5.29 (s, 1H), 4.25-4.15 (p, 4H), 4.2 (s, 1H), 3.73-3.64 (m, 8H), 3.58 (s, 4H) 1.45-1.30 (m, 23H). MS (ESI): m/z 887.0 $[(M+1)]^+$.

Compound 5-25: Compound **5-24** (4.52 g, 5.1mmol) was added into 40 ml of 1:1 DCM: TFA solution and stirred for an hour. After which the solvent was evaporated under reduced pressure to afford compound **5-25** and used without further purification. ^{1}H-NMR (300 MHz, solvent MS (ESI): m/z 787.3 $[(M+1)]^+$.

Compound 5-26: To a solution of **5-26** (1g, 1.27mmol) in dry DMF was added Biotin-NHS (0.43 g, 1.27mmol) and DIEA (0.66 ml, 3.81 mmol) and stirred for 12 h at RT. Upon removal of the solvent *in vacuo*, the resulting solution was taken into DCM and washed with saturated $NaHCO_3$ (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated in vacuo to afford a light brownish white solid which was further purified by flash chromatography (Silica gel, dichloromethane/ethanol = 9.5 : 0.5) to afford compound **5-26** (0.84 g, 66%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.87 (d, *J* = 2.54 Hz), 7.47 (d, *J* = 2.7 Hz), 7.43-7.36 (m, 2H), 7.29-7.34 (m, 2H), 7.55-7.22 (d, *J* = 3.0 Hz, 2H), 6.36 (broad, 2H), 5.94 (s, 1H), 5.78 (s, 1H), 4.32-4.11 (m, 6H), 3.61-3.58 (m, 8H), 3.28-3.20 (m, 2H), 3.01-2.87 (m, 2H), 2.01-2.06 (t, 2H), 1.69-1.41 (m, 6H), 1.36-1.19 (m, 12H). (ESI): m/z 1035.5 $[(M+23)]^+$.

Compound 5-27: Compound **5-26** (0.84g, 0.83 mmol) was added to 30 ml of 20 % piperidine in DMF and allowed to stir for 30 minutes, Upon concentration *in vacuo* followed by short flash chromatography purification (Silica gel, dichloromethane/ethanol = 8:2) afford compound **5-27** (0.62 g, 91%). MS (ESI): m/z 1012 $[(M+1)]^+$.

Compound 5-28: To a solution of compound **5-27** (0.62 g) in dry DMF was added Cy3-NHS (0.41g, 0.76mmol) and DIEA (0.39 ml, 2.28 mmol) and stirred for 12 h at RT, upon evaporation of the solvent in vacuo and purification by flash chromatography (silica gel, dichloromethane/ethanol = 9:1) afford compound **5-28** (0.31 g, 34%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.46-8.35 (t, 1H), 7.49 (d, *J* = 2.52 Hz, 2H), 7.31(d, *J* = 2.56 Hz), 7.23-7.14 (m, 8H), 6.40 (m, 2H), 5.83 (s, 1H), 5.67 (s, 1H), 5. 4.59-4.49 (m, 3H), 4.23-4.16 (p,

4H), 4.05 (broad, 1H), 4.09 (t, 3H), 3.63-3.58 (m, 8H), 3.27-3.15 (m, 4H, 2.30-2.24 (m, 2H), 1.37 (t, 6H).MS (ESI): m/z 1215.7 $[(M+1)]^+$.

Final probe 5-17: To a solution of compound **5-27** (0.2 g, 0.16mmol) in dry DCM was added TMSI (0.08 ml, 0.56 mmol) under ice cold condition and allowed to stir for 3.5 hrs at RT. After which the reaction was quenched with 10 % TEA in water (10 ml) and extracted with Chloroform (3 x 10 ml). The aqueous layer was separated and concentrated in vacuo to afford a red solid. Which was further purified by reverse phase HPLC to afford pure compound **5-17** (0.041 g, 22%) as a dark red solid. ^{1}H-NMR (300 MHz, CD_3OD) δ 8.56-8.47 (t, 1H), 7.52 (d, *J* = 2.54 Hz, 2H), 7.35(d, *J* = 2.54 Hz), 7.32-7.23 (m, 8H), 6.43 (d, *J* = 4.42 Hz, 2H), 5.81 (s, 1H), 5.65 (s, 1H), 5.0 (s, 1H) 4.43-4.47 (t, 2H), 4.24-4.26 (m, 1H), 4.12 (t, 3H), 3.45-3.43 (m, 8H), 3.16-3.09 (m, 4H), 2.92-2.90 (m, 1H), 2.87-2.84 (t, 3H), 2.64-2.68 (d, *J* = 4.3 Hz, 1H), 2.36-2.34 (m, 2H), 1.84-1.21 (m, 18H), 1.13-1.09 (m, 2H). (MS (ESI): m/z 1159.6 $[(M+1)]^+$.

7.6 Bioimaging using small molecule probes

7.6.1 Experimental Details of the Synthesis

7-Hydroxycoumarin-4-acetic acid (6-9): Resorcinol (5 g, 45 mmol) was dissolved in 50 ml of 70% H_2SO_4 under ice-cold conditions, and the solution was allowed to stir for 30 minutes. Acetone carboxylic acid (6.6 g, 45 mmol) was added in 5 portions. The mixture was allowed to stir further for 4 hrs, before pouring onto crushed ice pieces. The precipitate formed was washed with water, ethyl acetate and dried overnight under

reduced pressure to afford compound **6-9** as a pure white solid (9 g, 91%). ^{1}H NMR (300 MHz, DMSO) δ 7.53 (d, J = 9.1 Hz, 1H), 6.80 (dd, *J* = 8.7 & 2.1 Hz, 1H), 6.73 (d, *J* = 2.1 Hz, 1H), 6.22 (s, 1H), 3.82 (s, 2H). ^{13}C-NMR (60 MHz, DMSO) δ 170.56, 161.10, 158.16, 154.94, 150.05, 127.65, 126.58, 112.93, 111.90, 102.23, 39.30. MS (ESI): m/z 221.0 $[(M+1)]^{+}$

7-Acetoxycoumarin-4-acetic acid (6-10): To compound **6-9** (1 g, 4.53 mmol) was added 25 ml of acetic anhydride and 5 ml of pyridine, and the reaction was allowed to stir at RT for 30 minutes. Upon removal of solvent in vacuo, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated in vacuo to afford **10** (30% yield). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.71 (d, *J* = 8.9 Hz, 1H), 7.23 (d, *J* = 5.6 Hz, 1H), 7.24(s, 1H), 6.47(s, 1H), 2.94 (s, 2H), 2.33 (s, 3H). MS (ESI): m/z 263.0 $[(M+H)]^{+}$

Acetic acid 4-[9-benzyl sulfanylcarbonylmethyl-carbamoyl)-methyl]-2-oxo-2H-chromen-7-yl ester (6-1): To a solution of **6-10** (0.162 g, 0.62 mmol) in dry THF was added EDC (0.13 g, 0.68 mmol) and HOBt (0.104 g, 0.68 mmol). The reaction was stirred for 30 minutes under ice-cold conditions, followed by addition of **6-24** (0.135 g, 0.62 mmol) dissolved in minimal amount of DMF. The resulting solution was stirred further for 6 hours at RT. Upon removal of the solvent in vacuo, the resulting mixture was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated

in vacuo to afford a white solid, which was further purified by flash chromatography (silica gel, dichloromethane/ethanol = 9:1) to afford **6-1** (0.15 g, 56%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.67 (d, J = 9.0 Hz, 1H), 7.27-7.23 (m, 7H), 6.38 (s, 1H), 4.1 (s, 2H), 3.73 (s, 2H), 2.91 (s, 2H), 2.32 (s, 3H). MS (ESI): m/z 448.0 $[(M+Na)]^+$.

4-(3,6-diacetoxy-3H-xanthen-9-yl)-isophthalic acid (6-13): Resorcinol 6-11 (5.7 g, 52 mmol) was dissolved in 50 ml of CH_3SO_3H. To this solution, trimellitic anhydride (5 g, 26 mmol) was added, and the reaction was heated at 80-85 ºC for 12 hours. The highly viscous solution was cooled to RT then poured into 10 volumes of ice-cold water. The resulting precipitate was collected and dried in vacuo to give compound 6-12 as a crude yellow solid which was used without further purifications. Compound 6-12 was dissolved in 50 ml of acetic anhydride and 10 ml of pyridine. The solution was heated at 85 ºC for 15 minutes. The resulting solution was poured into ice-cold water to give the crude precipitate of 6-13, which, upon recystallization with DCM, afforded pure 6-13 as white crystals (5.17 g, 53%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.75 (s, 1H), 8.38 (d, J = 8.04 Hz, 1H), 7.29 (d, J = 8.04, 1H), 7.15-7.09 (m, 2H), 6.85-6.81 (m, 4H), 2.31 (s, 6H). ^{1}H-NMR (60 MHz, $CDCl_3$) δ 168.71, 167.46, 165.85, 155.48, 152.12, 150.69, 133.24, 132.47, 130.31, 129.23, 128.45, 126.14, 115.48, 110.41, 109.30, 81.21, 19.85. MS (ESI): m/z 461 $[(M+1)]^+$.

N-Benzylsulfanylcarbonylmethyl-6-(3,6-dihydroxy-3H-xanthen-9-yl)-isophthalamic acid (FL) (6-2): To a solution of **6-13** (0.424 g, 0.92 mmol) in dry THF was added EDC (0.176 g, 0.92 mmol) and HOBt (0.14 g, 0.92 mmol). The solution was stirred for 30

minutes under ice-cold conditions, followed by addition of **6-24** (0.2523 g, 1.16 mmol) dissolved in minimal amount of DMF. The resulting solution was stirred further for 6 hours at RT. Upon removal of the solvent in vacuo, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to afford a crude mixture, which was further purified by flash chromatography (silica gel, dichloromethane/methanol = 9:1) to afford **6-2** (0.32 g, 56%). ^{1}H -NMR (300 MHz, $CDCl_3$ δ 8.52 (s, 1H), 8.22 (d, *J* = 8.04 Hz, 1H), 7.25-7.21 (m, 6H), 7.12-7.09 (m, 2H), 6.81-6.75 (m, 4H), 4.39 (d, *J* = 5.61 Hz, 2H), 4.15 (s, 2H), 2.3 (s, 6H). ^{13}C-NMR (60 MHz, DMSO) δ 196.52, 168.78, 168.37, 165.84, 155.31, 152.20, 151.42, 136.72, 134.08, 129.58, 127.74, 127.50, 127.33, 126.45,125.61, 124.50, 123.91, 115.53, 110.46, 82.09, 49.48, 32.90, 20.98. MS (ESI): m/z 623.9 $[(M+1)]^+$.

5-Carboxy-tetramethylrhodamine (6-15): Trimellitic anhydride (1.00 g, 5.2 mmol) and 3-dimethylaminophenol, **6-14** (0.72 g, 10.5 mmol) was refluxed in toluene (50 ml) for 12 hours. Upon cooling to RT, the resulting precipitate was collected by filtration, and further purified by flash chromatography (silica gel, dichloromethane/methanol/acetic acid = 8:1.9:0.1) to afford pure **6-15** as a dark purple solid (0.811 g, 36%). ^{1}H -NMR (300 MHz, $CDCl_3$) δ 8.52 (s, 1H), 8.12 (d, *J* = 8.61 Hz, 1H), 7.59 (d, *J* = 7.62 Hz, 1H), 6.82 (m, 2H), 6.51 (d, J = 5.61 Hz, 2H). 6.09 (s, 2H), 3.28 (s, 12H). MS (ESI): m/z 431.2 $[(M+1)]^+$

Amino thioacetic acid S-benzyl ester of 5-carboxy-tetramethylrhodamine (TMR) (6-3): To a solution of **6-15** (0.05 g, 0.116 mmol) in dry THF was added EDC (0.022 g, 0.116 mmol) and HOBt (0.018 g, 0.116mmol). The solution was stirred for 30 minutes under ice-cold conditions, followed by addition of **6-24** (0.025 g, 0.116 mmol) dissolved in minimal amount of DMF. The resulting solution was stirred further for 6 hours at RT. Upon removal of the solvent in vacuo, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to afford a crude mixture, which was further purified by flash chromatography (silica gel, dichloromethane/ethanol = 9:1) to afford **6-3** (0.024 g, 35%) as a dark purple solid. ^{1}H - NMR (300 MHz, CD_3OD) δ 8.78 (s, 1H), 8.27 (d, *J* = 8.01 Hz, 1H), 7.55 (d, *J* = 7.83 Hz, 1H), 7.28 (d, *J* = 6.45, 2H), 7.20-7.26 (m, 5H), 7.09-6.99 (m, 4H), 4.40 (s, 2H), 4.15 (s, 2H), 3.30 (s, 12H). MS (ESI): m/z 594.2 $[(M+1)]^+$.

Acetoxynapthofluorescein (6-18): Trimellitic anhydride (1.93 g, 10 mmol) was dissolved in methanesulfonic acid (1 M). **6-16** (3.2 g, 20 mmol) was added and the solution was heated at 100 ^{o}C for 12 hrs. Upon cooling to RT, the solution was poured into 8 volumes of ice-cold water. The resulting red precipitate was collected and dried in vacuo to give crude **6-17**, which was used directly without further purifications. MS (ESI): m/z 477.4 $[(M+1)]^+$. Crude **6-17** was dissolved in 50 ml of acetic anhydride and 10 ml of pyridine. The resulting mixture was heated at 80-90 ^{0}C for 15 minutes**.** Upon concentration to dryness, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 50 ml), water (2 x 50 ml) and brine (2 x 50 ml). The organic layer was dried

over anhydrous Na_2SO_4, and concentrated in vacuo to give crude **6-18** (46% yield for 2 steps based on the crude product), which was confirmed by MS ((ESI): m/z 561.0 $[(M+1)]^+$ and used without further purifications.

Amino thioacetic acid S-benzyl ester of acetoxynapthofluorescein (CF) (6-4): To a THF solution of crude **6-18** (0.065 g, 0.116 mmol), obtained from above reaction, was added EDC (0.022 g, 0.116 mmol) and HOBt (0.018 g, 0.116 mmol). The solution was stirred for 30 minutes under ice-cold conditions, followed by addition of **6-24** (0.025 g, 0.116 mmol) dissolved in minimal amount of DMF. The resulting solution was stirred further for 6 hours at RT. Upon removal of the solvent in vacuo, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to afford a crude mixture, which was further purified by reverse phase HPLC to afford pure **6-4** (0.025 g, 30% yield). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.62 (s, 1H), 8.22(d, *J* = 6.42 Hz, 1H), 8.14-8.11 (m, 2H), 7.64 (s, 2H), 7.48 (d, *J* = 8.03 Hz, 1H), 7.31-7.27 (m, 9H), 7.19 (d, *J* = 7.62 Hz, 2H), 4.2 (d, *J* = 6.0 Hz, 2H), 4.01 (s, 2H), 2.37 (s, 6H). MS (ESI): m/z 724.1 $[(M+1)]^+$

5-(2-Oxo-hexahydro-thieno[3,4-d]imidazol-6-yl)-pentanoylamino]-thioacetic acid S-benzyl ester (6-5): Biotin **6-19** (0.156 g, 0.64 mmol) was dissolved in 40 mL of dry THF. To this solution was added EDC (0.123 g, 0.64 mmol) and HOBt (0.098 g, 0.64 mmol) under ice-cold conditions. The reaction was allowed to stir for 30 minutes, followed by addition of **6-24** (0.139 g, 0.64 mmol) dissolved in minimal amount of DMF. The resulting solution was allowed to stir overnight at RT. Upon removal of the solvent

in vacuo, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to afford **6-5** (0.110 g, 42%). ^{1}H-NMR (300 MHz, DMSO) δ 7.31-7.28 (m, 5H), 6.40-6.35 (broad, 2H), 4.37-4.32 (m, 2H), 4.22 (d, *J* = 5.61 Hz, 2H), 4.09 (s, 2H), 3.1-3.0 (m, 1H), 2.88-2.78 (m, 2H), 2.19-2.14 (m, 2H), 1.60-1.48 (m, 6H). MS (ESI): m/z 408.0 $[(M+1)]^+$.

(4-Benzoyl-benzoylamino)-thioacetic acid S-benzyl ester (6-6): To a solution of **6-20** (0.136 g, 0.6 mmol) in dry THF (40 ml) was added EDC (0.115 g, 0.6 mmol) and HOBt (0.092 g, 0.6 mmol). The solution was stirred for 30 minutes under ice-cold conditions, followed by addition of **6-24** (0.136 g, 0.6 mmol) dissolved in minimal amount of DMF. The resulting solution was stirred further for 6 hours at RT. Upon removal of the solvent in vacuo, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 ml), water (2 x 30 ml) and brine (2 x 30 m). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to afford a white solid, which was further purified by flash chromatography (silica gel, dichloromethane/ethanol = 9.5:0.5) to afford **6-6** (0.157 g, 68%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.94-7.91 (m, 2H), 7.76-7.73 (m, 4H), 7.61-7.56 (m, 1H), 7.48-7.43 (m, 2H), 7.27-7.20 (m, 5H), 4.37 (d, *J* = 5.64 Hz, 2H), 4.12 (s, 2H). ^{13}C-NMR (60 MHz, $CDCl_3$) 195.88, 194.66, 166.68, 140.39, 138.69, 136.66, 134.20, 132.89, 131.82, 131.17, 130.05, 129.99, 128.99, 128.80, 128.23, 128.16, 127.45, 127.38, 127.10, 49.39, 33.00. MS (ESI): m/z 389.47 $[(M+1)]^+$.

2-Nitrobenzyl tosylate (6-22): 2-Nitrobenzyl alcohol (1 g, 6.5 mmol) was dissolved in 40 ml of dry DCM and 1.36 ml of triethylamine. p-Toluenesulfonyl chloride (1.24 g, 6.5 mmol) was added and the solution was stirred for 2 hours before quenching with water. The organic layer was separated and washed with 1 N HCl(1 x 50 ml), water (1 x 50 ml), 2 N NaOH (1 x 50 ml), water (1 x 50 ml) and brine (1 x 50 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to give a dull white solid, which was further purified by flash chromatography (silica gel, dichloromethane/ethanol = 9.5:0.5) to give 2-nitrobenzyl tosylate, **6-22** (1.57 g, 79%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.08-7.99 (dd, *J* = 9.21 & 8.04 Hz, 1H), 7.81 (d, *J* = 8.4 Hz, 1H), 7.72-7.62 (m, 3H), 7.50 (d, *J* = 8.85 Hz, 1H) 7.33 (d, *J* = 8.43 Hz, 2H), 4.93 (s, 2H), 2.42 (s, 3H). ^{13}C-NMR (60 MHz, $CDCl_3$) δ 149.33, 146.12, 137.71, 135.27, 134.11, 133.09, 132.69, 132.57, 128.12, 123.89, 42.72, 20.64. MS (ESI): m/z 329.9 $[(M+Na)]^+$.

"Caged" Fluorescein (C2FL) (6-23): Crude **6-12** was first purified by recrystallization. The recrystallized product (0.602 g, 1.6 mmol) was dissolved in 30 ml of DMF containing K_2CO_3 (1.1 g, 5 mmol). 2-nitrobenzyl tosylate (0.492 g, 1.6 mmol) was added to the solution and the reaction was heated at 60 ºC for 6 hours in a RBF covered with Aluminium foil to maintain dark conditions. After 6 hrs, the solvent was removed in vacuo and the residual mixture was washed with 1 N HCl, water and brine followed by drying in anhydrous Na_2SO_4 and concentration under reduced pressure to afford **6-23** (0.64 g, 62%). The work up was strictly done in dark. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.44 (s, 1H), 8.16 (d, *J* = 8.11 Hz, 1H), 8.00-7.97 (m, 2H), 7.75-7.85 (m, 2H), 7.60-7.56 (m,

4H), 7.45 (d, *J* = 8.10 Hz, 1H), 7.35 (d, *J* = 7.86 Hz, 2H), 7.22-7.10 (m, 4H), 5.51 (s, 4H). MS (ESI): m/z 647.0 $[(M+1)]^+$.

Amino thioacetic acid S-benzyl ester of C2FL (6-7): The reaction and workups were carried out in dark. To a solution of **6-23** (0.084 g, 0.13 mmol) in dry THF was added EDC (0.027 g, 0.143 mmol) and HOBt (0.022 g, 0.143 mmol). The solution was stirred for 30 minutes under ice-cold conditions, followed by addition of **6-24** (0.028 g, 0.13 mmol) dissolved in minimal amount of DMF. The resulting solution was stirred further for 6 hours at RT. Upon removal of the solvent in vacuo, the residue was taken into ethyl acetate and washed with 1 N HCl (2 x 30 mm), water (2 x 30 ml) and brine (2 x 30 ml). The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to afford a crude mixture, which was further purified by flash chromatography (silica gel, dichloromethane/ethanol = 9.8:0.2) to afford **6-7** (0.040 g, 38%). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 8.38 (s, 1H), 8.29 (d, *J* = 8.13 Hz, 1H), 8.08-8.11 (m, 2H), 7.99-7.90 (m, 2H), 7.71-7.68 (m, 4H), 7.65 (d, J = 8.10 Hz, 1H), 7.60-7.49 (m, 6H), 7.25 (s, 5H), 5.29 (s, 4H), 4.68 (s, 2H), 4.19 (s, 2H). MS (ESI): m/z 810.1 $[(M+1)]^+$.

7.6.2 Fluorescence and *in vitro* labeling experiments

10 µl each of the stock solutions of probes **6-1**, **6-2**, **6-3** and **6-4** (1 mM in DMSO) was added to 100 µl of 1 mM K_2CO_3 solution, and the mixture was allowed to stand at RT for > 5 min for the deacylation reaction to occur, resulting in the release of very strong fluorescence in the solution. 10 µl of the solution was subsequently diluted with 100 µl of distilled water, and the resulting solution was transferred to a black 96-well microtitre plate (Nunc, USA), where the excitation and the emission spectra of the probes

were recorded using a SpectraMax™ GeminiXS fluorescence microplate reader (Molecular Devices, USA).

In vitro Protein labeling procedures With N-terminal cysteine proteins

Cloning and expression of N-terminal cysteine proteins will be reported elsewhere. Purified enhanced green fluorescent protein (EGFP) having an N-terminal cysteine was incubated for up to 24 hours with probes **6-2, 6-3, 6-4** & **6-5**, respectively, in order to assess our labeling strategy. Probes were prepared as 200 μM stocks (25 x in DMSO) and stored at -20 ^{o}C. In a typical labeling reaction, 6 μl of each probe (final concentration: 8 μM) was added to 50 μl of pure protein (1 mg/ml) dissolved in 1 x PBS (final concentration of protein: ~2 nM), with or without 1 mM DTT, and the reaction was topped up with 1 x PBS to a final volume of 150 μl. DTT was shown in our experiments, where live cells were used, to reduce background labelings. At specific time intervals, 15 μl of the reaction was withdrawn and quenched by addition of 1.7 μl of 100 mM cysteine to the reaction mixture (final concentration of cysteine: 10 mM), followed by denaturation with SDS-PAGE loading dye at 95 ^{o}C for 3 min. The loading dye also serves to hydrolyze the diacetates groups on probes **6-2** and **6-4**, thereby releasing the fluorescence. Upon separation with a 12% SDS-PAGE gel, the fluorescence labeling of EGFP by probes **6-2**, **6-3** and **6-4** was conveniently visualized by scanning the resulting gel with a Typhoon™ 9200 fluorescence scanner (Amersham Biosciences, USA). Fluorescence intensity of each protein band was analysed using the software, Image Quant 5.2, preinstalled on the instrument. In all cases, more than 75% of labeling was obtained in 3 hours for all three probes. For the labeling of EGFP with probe **6-5**, an anti-biotin Western blot was performed to visualize the amount of biotinylated EGFP. Briefly,

following SDS-PAGE separation, the resulting gel was electroblotted onto a polyvinylidene difluoride (PVDF) membrane (BioRad, USA) and blocked for 1 h with 5% non-fat dry milk in PBST (phosphate buffered saline, pH 7.4 with 0.1% Tween 20) (**Fig. 7.18**) The membrane was incubated with anti biotin-conjugated HRP (NEB, USA) in an 1:3000 dilution in milk-PBST for 1 h and then washed with PBST (3 x 15 min). Visualization was done with the Enhanced ChemiLuminescent ECL™ kit (Amersham).

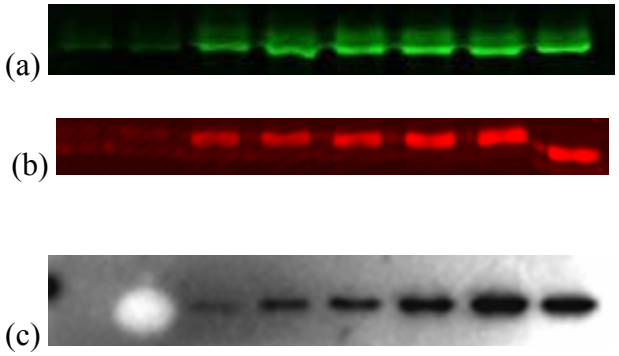


Fig. 7.20 SDS-PAGE of purified N-terminal cysteine EGFP labeled with a) probe **6-2** and b) probe **6-3** followed by visualization with fluorescence gel scanner; and c) probe **6-5** followed by Western blot with anti biotin- conjugated HRP.

7.6.3 Protein labeling inside live bacterial cells.

A DNA construct containing N-terminal Cysteine GST was transformed into the E.coli expression strain ER2566 (NEB) and grown in 100mg/L ampicillin containing LB media at 37°C in a 250rpm air shaker. At OD_{600} ~ 0.6, protein expression was induced with 0.3mM IPTG (isopropyl-β-D-thiogalactoside). Cells were grown for 12h at room temperature to optimize protein expression as well as self-cleavage of the intein-fusion protein to generate N-terminal cysteine proteins in vivo. For labeling, 8μM of the TMR probe (probe **6-3**) was added directly to the LB media and the cells incubated for 24h.

Cells were harvested by centrifugation at 4000rpm for 10min and resuspended in 1x PBS buffer (pH 7.4) containing 10% glycerol and left standing for 30min. This procedure was repeated 3 times for complete removal of any free probe. Subsequently, cells were used for either fluorescence microscope or SDS-PAGE experiments.

For fluorescence microscope experiments, cells were mounted on clean glass slides coated with 1.5% agarose. Fluorescence images were recorded with a AxioSkop 40

fluorescent microscope (Carl Zeiss, Germany) equipped with a cooled CCD camera (AxioCam, Zeiss) using a 100x oil objective. TMR fluorescence imaging was recorded using a BP546/12 excitation and LP590 emission filter (Zeiss).

For SDS-PAGE experiments, the labeled cells, upon washes as mentioned above, were harvested by centrifugation. The cell pellet was treated directly with SDS-PAGE loading dye, and analyzed by 12% SDS-PAGE gel and fluorescence scanning (Typhoon 9200, Amersham Biosciences). The results are given in **Fig. 6.4.**

Chapter 8

Concluding Remarks

High-throughput enzymology occupies a pivotal role in the modern drug discovery programs and relies on high-throughput amenable chemical reactions that allow rapid synthesis of diverse chemical libraries for the interrogation of different classes of enzymes. We have successfully applied 'click' chemistry, solid-phase 'amide-bond' formation reaction and Microwave-assisted synthesis as high-throughput platforms to synthesize large and diverse libraries of inhibitors targeting various PTPs as well as other vital enzymes. The highlight of all the above described approaches is their simplicity and the high-quality of the final products. The second part of my thesis demonstrates our successful synthesis of a novel class of activity-based probes targeting various proteases. Further, these probes were used to generate unique substrate fingerprint profiles of proteases which could answer to potential problems like substrate identification, catalytic mechanism of unknown proteases. The last part of my thesis narrates our successful synthesis of a set of small molecule probes for the site-specific labeling of N-terminal cysteine proteins in live cells. I believe that all the above approaches will be valuable assets in the tool box of "Catalomics".

Chapter 9

References

1. Venter, J.C. et al. *Science* **2001**, *291*, 1304.
2. Drews, J. *Science* **2000**, *287*, 1960.
3. Sun, H.; Chattopadhaya, S.; Wang, J.; Yao, S. Q. *Anal. Bioanal. Chem.* **2006**, *386*, 416.
4. (a) Hunter, T. *Cell* **2000**, *100*, 113. (b) Guo, X.-L.; Shen, K.; Wang, F.; Lawrence, D. S.; Zhang, Z.-Y. *J. Biol. Chem.* **2002**, *277*, 41014.
5. Bialy, L.; Waldmann, H. *Angew. Chem., Int. Ed.* **2005**, *44*, 3814.
6. Zhang, Z.-Y. *Curr. Opin. Chem. Biol.* **2001**, *5*, 416.
7. (a) Sharpless, K. B.; Manetsch, R. *Exp. Opin. Drug Discovery* **2006**, *1*, 525. (b) Kolb, H. C.; Sharpless, K. B. *Drug Disco*V*ery Today* **2003**, *8*, 1128.
8. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596.
9. Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.*, **2002**, 67, 3057.
10. Wu, P.; Fokin, V. V. *Aldrich. Acta.* **2007**, *40,* 7.
11. Arkin, M. R.; Wells, J. A. *Nat. Rev. Drug Discovery* **2004**, *3*, 301.
12. Szczepankiewicz, B. G. et al. *J. Am. Chem. Soc.* **2003**, *125*, 4087.
13. Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9367.

14. Birk, A.; Lin, Y.-C.; Elder, J.;Wong, C.-H. *Chem. Biol.* **2002**, *9*, 891.

15. Srinivasan, R.; Uttamchandani, M.; Yao, S. Q. *Org. Lett.* **2006**, *8*, 713.

16. Xie, J.; Seto, C.T. *Bioorg. Med. Chem*. **2007**, *15*, 458 – 473.

17. Wang, J.; Uttamchandani, M.; Li, J.; Hu, M.; Yao, S. Q. *Org. Lett*. **2006**, *8*, 3821.

18. Mocharla, V.P.; Colasson, B.; Lee, L.V.; Roper, S.; Sharpless, K. B.;Wong, C. – H.; Kolb, H.C. *Angew. Chem. Intl. Ed*. **2005**, *44*, 116.

19. Sears, P.; Wong, C, -H. *Cell. Mol. Life Sci.* **1998**, *54*, 223.

20. Lee, L.V.; Mitchel, M.L.; Huang, S. J.; Fokin, V.V.; Sharpless, K.B.; Wong, C. – H. J. Am. Chem. Soc. **2003**, *125*, 9588.

21. Wu, C.-Y. et al. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10012.

22. Krasinski, A.; Radic, Z.; Manetsch, R.; Raushel, J.; Taylor, P.; Sharpless, K. B.; Kolb, H. C, *J. Am. Chem. Soc*. **2005**, *127*, 6686.

23. Ng, S.L., Yang, P.-Y.; Chen, K.Y.-T.; Srinivasan, R.; Yao, S.Q, *Org. Biomol. Chem.* **2008**, *6*, 844.

24. Yang, P-Y.; Wu, H.; Lee, M.Y.; Xu, A.; Srinivasan, R.; Yao, S.Q, *Org. Lett.*, **2008**, *10*, 1881.

25. Brik, A.; Lin, Y.-C.; Elder, J.; Wong, C.-H. *Chem. Biol.* **2002**, *9*, 891.

26. Lee, S.-G.; Chmielewski, J. *Chem. Biol.* **2006**, *13*, 421.

27. Best, M. D.; Brik, A.; Chapman, E.; Lee, L. V.; Wong, C.-H. *ChemBioChem* **2004**, *5*, 811.

28. Wu, C.-Y.; Chang, C.-F.; Chen, J. S.-Y.; Lee, S.-T.; Wong, C.-H.; Lin, C.-H. *Angew. Chem. Int. Ed.* **2003**, *42*, 4661.

29. Wu, C.-Y.; King, K.-Y.; Kuo, C.-J.; Fang, J.-M.; Wu, Y.-T.; Ho, M.-Y.; Liao, C.-L.; Shie, J.-J.; Liang, P.-H.; Wong, C.-H. *Chem. Biol.* **2006**, *13*, 261.

30. R. B. Merrifield, *J. Am. Chem. Soc*. **1963**, *85*, 2149.

31. Bunin, B.A.; Ellman, J. A. *J. Am. Chem. Soc*. **1992**, *114*; 10997.

32. Furka, A.; Sebestyen, F.; Asgedom, M.; Dibo, G. *Int. J. Pept. Protein Res*. **1991**, *37*, 487

33. Geysen, H. M.; Meloen, R. H.; Bartteling, S. J. *Proc. Natl. Acad. Sci, U.S.A*, **1984**, *81*, 3998.

34. Plunkett, M. J.; Ellman, J. A. *J. Org. Chem*. **1995**, 60, 6006.

35. Fivush, A. M.; Willson, T. M. *Tetrahedron Lett.* **1997**, *38*, 7151.

36. Kappe, C. O.; Dallinger. D. *Nature Rev. Drug. Discov*. **2006**, *5*, 51.

37. Pisani, L.; Prokopcova, H.; Kremsner, J. M.; Kappe, C. O, *J. Comb. Chem*. **2007**; *9*, 415.

38. Perez, R.; Beryozkina, T.; Zbruyev, O. I.; Haas, W.; Kappe, C. O, *J. Comb. Chem.* **2002**, 4, 501.

39. Evans, M. J.; Cravatt, B. F. *Chem. Rev*. **2006**; *106*, 3279.

40. Liau, M. L. ; Panicker, R. C. ; Yao, S.Q. *Tetrahedron Lett.* **2003**, *44*, 1043.

41. Wang, G.; Uttamchandani, M. ; Chen, G.Y.J. ; Yao, S.Q. *Org. Lett.* **2003** *5*, 737.

42. Meng, L.; Kwok, B. H.; Sin, N.; Crews, C. M. *Cancer Res.* **1999**, *59*, 2798.

43. Liu, Y.; Patricelli, M. P.; Cravatt, B. F. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14694

44. Kumar, S.; Zhou, B.; Liang, F.; Wang, W. Q.; Huang, Z.; Zhang, Z. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7943.

45. Evans, M. J.; Saghatelian, A.; Sorenson, E. J.; Cravatt, B. F. *Nat. Biotechnol*., **2005,** 23, 1303.

46. Wright, A. T.; Cravatt, B. F. *Chem. Biol*., **2007**, 14, 1043.

47. Wang, J.; Uttamchandani, M.; Li, J.; Hu, M.; Yao, S. Q. *Chem. Commun.*, **2006**, 3783.

48. Blum, G.; Degenfeld, G.; Merchant, M. J.; Blau, H. M.; Bogyo, M. *Nat. Chem. Biol*., **2007**, 3, 668.

49. Uttamchandani, M.; Walsh, D.P.; Yao, S.Q.; Chang, Y.T. *Curr. Opin. Chem. Biol.* **2005**, *9*, 4.

50. Hu, Y.; Chen, G.Y.J.; Yao, S.Q, *Angew. Chem. Intl. Ed.* **2005**, *44*, 1048.

51. Michalet, X.; Kapanidis, A. N.; Laurence, T.; Pinaud, F.; Doose, S.; Pflughoefft, M.; Weiss, S. *Ann. Rev. Biophy. Biomol*. **2003**, *32*, 161.

52. Miyawaki, A.; Sawano, A.; Kogure, T. *Nat. Cell Biol.* **2003**, S1.

53. Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y. *Nat. Rev. Mol. Cell Biol*. **2002**, *3*, 906.

54. Tsien, R. Y. *Ann. Rev. Biochem*. **1998**, *57*, 509.

55. Griffin, B. A.; Adams, S. R.; Tsein, R. Y. *Science* **1998**, *28*, 269.

56. Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. *Nat. Biotechnol*. **2003**, *21*, 86.

57. Pegg, A. E. *Mutat. Res*. **2000**, *462*, 83.

58. Guignet, E. G.; Hovius, R.; Vogel, H. *Nat. Biotechnol*. **2004**, *22*, 440.

59. Kapanidis, A. N.; Ebright, Y. W.; Ebright, R. H. *J. Am. Chem. Soc.* **2001**, *123*, 12123.

60. Yeo, D. S. Y.; Srinivasan, R.; Uttamchandani, M.; Chen, G. Y. J.; Zhu, Q.; Yao, S. Q. *Chem. Commun.* **2003**, 2870.

61. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776.

62. Hajduk, J. P.; Greer, J. *Nat. Rev. Drug. Discov*., **2007**, 5, 211.

63. Erlanson, D. A.; McDowell, R. S.; He, M. M.; Randal, M.; Simmons, R. L.; Kung, J.; Waight, A.; Hansen, S. K. *J. Am. Chem. Soc.*, **2003**, 125, 5602.

64. Soellner, M. B.; Rawls, K. A.; Grundner, C.; Alber, T.; Ellman, J. A. *J. Am. Chem. Soc.*, **2007**, 129, 9613.

65. Sharpless, K. B.; Manetsch, R. *Exp. Opin. Drug Discovery* **2006**, *1*, 525.

66. Kolb, H. C.; Sharpless, K. B. *Drug Discovery Today* **2003**, *8*, 1128.

67. Srinivasan; R.; Li, J.; Ng, S.L.; Kalesh, K.A.; Yao, S.Q, *Nat. Protocols*, **2007**, *2*, 2655.

68. Johnson, T. O.; Ermolieff, J.; Jirousek, M. R. *Nat. Rev. Drug Discovery* **2002**, *1*, 696.

69. Puius, Y. A.; Zhao, Y.; Sullivan, M.; Lawrence, D. S.; Almo, S. C.; Zhang, Z.-Y. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 13420.

70. Liu, G.; et al. *J. Med. Chem.* **2003**, *46*, 4232–4235.

71. Liu, G.; et al. *J. Med. Chem.* **2003**, *46*, 3437– 3440.

72. Szczepankiewicz, B. G.; Liu, G.; Hajduk, P. J.; Abad-Zapatero, C.; Pei, Z.; Xin, Z.; Lubben, T. H.; Trevillyan, J. M.; Stashko, M. A.; Ballaron, S. J.; Liang, H.; Huang, F.; Hutchins, C. W.; Fesik, S. W.; Jirousek, M. R. *J. Am. Chem. Soc.* **2003,** 125, 4087.

73. Zhu, Q., Li, D.B., Uttamchandani, M., Yao, S.Q. *Bioorg. Med. Chem. Lett.* **2003** *13*, 1033.

74. Bräse, S.; Gil, C.; Knepper, K.; Zimmermann, V. *Angew. Chem. Int. Ed.,* **2005**, *44*, 5188.

75. Golebiowski, A.; Klopfenstein, S. R.; Portlock, D. E. *Curr. Opin. Chem. Biol.* **2003**, *7*, 308.

76. Hu, Y.; Uttamchandani, M.; Yao, S. Q. *Comb. Chem. High Throughput Screen.* **2006**, *9*, 203.

77. Wu, C.-Y.; Chang, C.-F.; Chen, J. S.- Y.; Lee, S.-T.; Wong, C.-H.; Lin, C.-H. *Angew. Chem., Int. Ed.* **2003**, *42*, 4661.

78. Brik, A.; Wu, C.-Y.; Wong, C.-H. *Org. Biomol. Chem.* **2006**, *4*, 1446.

79. Wu, C.-Y.; King, K.-Y.; Kuo, C.-J.; Fang, J.-M.; Wu, Y.-T.; Ho, M.-Y.; Liao, C.-L.; Shie, J.-J.; Liang, P.-H.; Wong, C.-H. *Chem. Biol.* **2006**, *13*, 261.

80. Raghavan, S.; Yang, Z.; Mosley, R. T.; Scheleif, W. A.; Gabryelski, L.; Olsen, D. B.; Stahlhut, M.; Kuo, L.C.; Emini, E. A.; Chapman, K. T.; Tata, J. R. *Bioorg. Med. Chem. Lett***., 2002**, 12, 2855.

81. Srinivasan; R.; Tan, L.P.; Wu, H.; Yao, S.Q. *Org. Lett.* **2008**, *10*, 2295.

82. Knepper, K.; Lormann, M. E. P.; Bräse, S. *J. Comb. Chem.,* **2004**, 6, 460

83. Gibson, C. L.; La Rosa, S.; Suckling, C. J. *Tetrahedron Lett.,* **2003**, 44, 1267.

84. Kappe, C.O. *Angew. Chem. Int. Ed.* **2004**, *43*, 6250.

85. Pothukanuri, S.; Winssinger, N. *Org. Lett*. **2007**, *9*, 2223.

86. Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. *Org. Lett*. **2000,** *2,* 1939.

87. Appukuttan, P.; Dehaen,W.; Fokin, V.V.; Van der Eycken, E. *Org Lett*. **2004**, 6, 4223.

88. Feldman, A. K.; Colasson, B.; Fokin, V. V. *Org. Lett.*, **2004**; *6*, 3897.

89. Q. Liu and Y. Tor, *Org. Lett.* 2003, **5**, 2571.

90. H. S. G. Beckmann and V. Wittmann, *Org. Lett.* **2007**, *9*, 1.

91. E. D. Goddard-Borger and R. V. Stick, *Org. Lett.* **2007**, *9*, 3797.

92. Iranpoor, N.; Firouzabadi, H..; Akhlaghinia, B.; Nowrouzi, N. *Tet lett*. **2004**, *45*, 3291.

93. Ju, Y.; kumar, D.; Varma, R.S. *JOC*, **2006**, *71*, 6697.

94. Park, S.H. *Bull. Korean Chem. Soc*. **2003**, *24*, 253.

95. Goddard, J. P.; Reymond, J.-L. *Curr. Opin. Biotechnol.* **2004**, *15*, 314.

96. Abecassis, V.; Pompon, D.; Truan, G. *Nucleic Acids Res.* **2000**, *28*, e88.

97. Toepert, F.; Knaute, T.; Guffler, S.; Pires, J.-R.; Matzdorf, T.; Oschkinat, H.; Schneider-Mergener, J. *Angew. Chem., Int. Ed.* **2003**, *10*, 1136-1140.

98. Grognux, J.; Reymond, J.-L. *ChemBioChem* **2004**, *5*, 826-831.

99. Chen, G. Y. J.; Uttamchandani, M.; Zhu, Q.; Wang, G.; Yao, S. Q. *ChemBioChem* **2003**, *4*, 336.

100. Greenbaum, D. C.; Arnold, W. D.; Lu, F.; Hayrapetian, L.; Baruch, A.; Krumrine, J. ; Toba, S.; Chehade, K.; Bromme, D.; Kuntz, I. D.; Bogyo, M. *Chem Biol.* **2002,** 9, 1085.

101. Powers, J. C.; Asgian, J. L.; Ekici, O. D.; James, K. E. *Chem. Rev*. **2002**; *102*; 4639.

102. Saghatelian, A.; Cravatt, B.F. *Curr. Opin. Chem. Biol.* **2005**, *9*, 62.

103. Hemelaar, J.; Galardy, P. J.; Borodovsky, A.; Kessler, B. M.; Ploegh, H. L.; Ovaa, H. J. Proteome Res. **2004**; *3*, 268

104. Zhu, Q.; Girish, A.; Chattopadhaya, S.; Yao, S. Q. *Chem. Commun.* **2004**, 1512.

105. Kidd, D.; Liu, Y.; Cravatt, B. F. *Biochemistry,* **2001**, 40, 4005.

106. Wang, G., Uttamchandani, M., Chen, G.Y.J., Yao, S.Q. *Org. Lett.* **2003**, *5*, 737.

107. Eisen, M. B.; Spellman, P. T.; Brown, P. O. Botstein, D. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14863.

108. Lo, L. C.; Pang, T. L.; Kuo, C. H.; Chiang, Y. L.; Wang, H. Y.; Lin, J. J. *J. Proteome Res*. **2002**, *1*, 35.

109. Zhu, Q.; Huang, X.; Chen, G. Y. J.; Yao, S. Q. *Tetrahedron Lett.* **2003**, *44*, 2669.

110. Chattopadhaya, S.; Srinivasan. R.; Yeo, S. Y. D.; Chen, G. Y. J.; Yao, S. Q. *Bioorg. Med. Chem*. **2009**, 17, 981.

111. Xu, M. Q.; Evans, T. C. *Methods* **2001**, *24*, 257.

112. Baker, R.T. *Curr. Opin. Biotechnol.*, **1996**, *7*, 541

113. Yeo, S.Y.D., Srinivasan, R.; Chen, G.Y.J.; Yao, S.Q. *Chem. Eur. J*. **2004**, *10*, 4664.

114. Chattopadhaya, S.; Bakar, F.B.A.; Srinivasan; R.; Yao, S.Q. *ChemBioChem*, **2008**, *9*, 677.

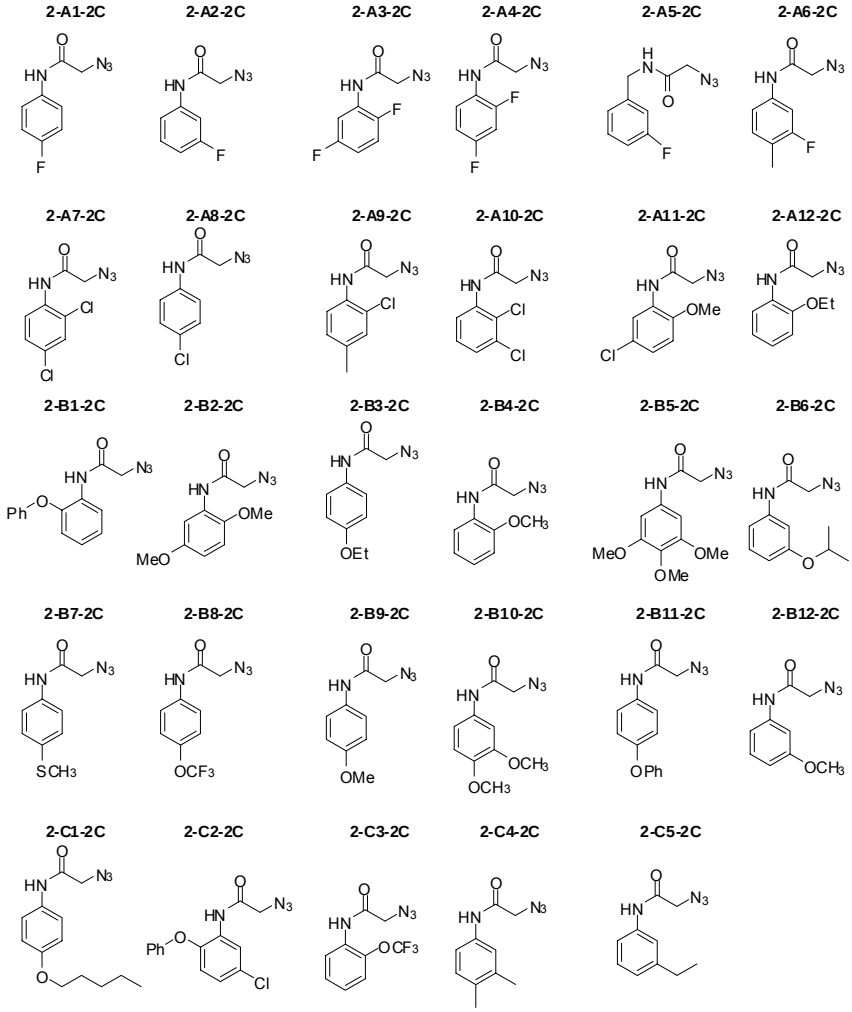
115. R. P. Haugland, *Handbook of fluorescent probes and research products*; 9th Edn., Molecular Probes: Eugene, OR, **2002**.

116. Schwartz, J. P.; Patterson, G. H. *Science*, **2003**, 300, 87.

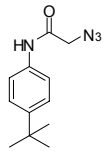
Chapter 10

Appendices

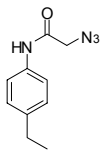
10.1 Structure of the azides synthesized by traceless solid-phase approach



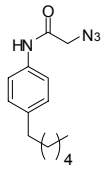
2-C6-2C



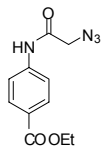
2-C7-2C



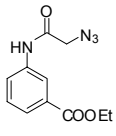
2-C8-2C



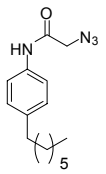
2-C9-2C



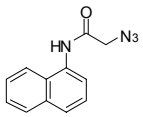
2-C10-2C



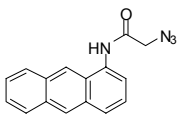
2-C11-2C



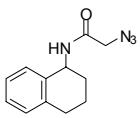
2-C12-2C



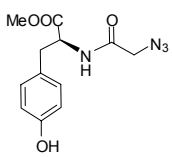
2-D1-2C



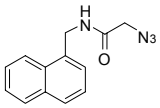
2-D2-2C



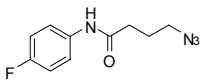
2-D3-2C



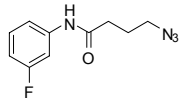
2-D4-2C



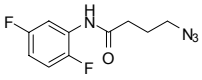
2-E1-4C



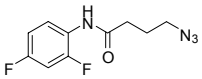
2-E2-4C



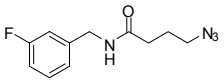
2-E3-4C



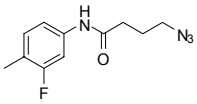
2-E4-4C



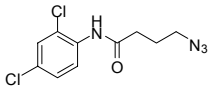
2-E5-4C



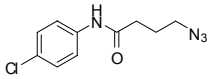
2-E6-4C



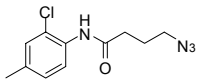
2-E7-4C



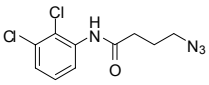
2-E8-4C



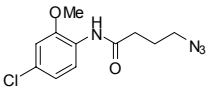
2-E9-4C



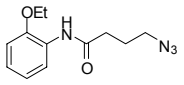
2-E10-4C



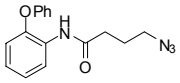
2-E11-4C



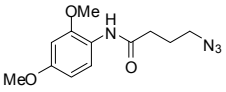
2-E12-4C



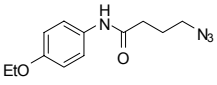
2-F1-4C



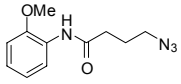
2-F2-4C



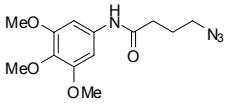
2-F3-4C



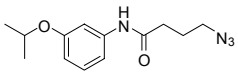
2-F4-4C



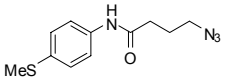
2-F5-4C



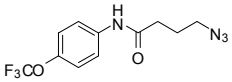
2-F6-4C



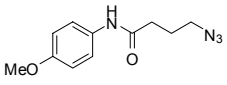
2-F7-4C



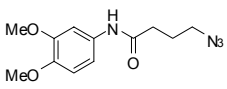
2-F8-4C



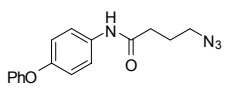
2-F9-4C



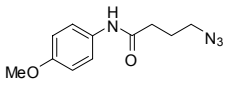
2-F10-4C



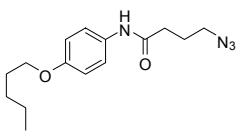
2-F11-4C



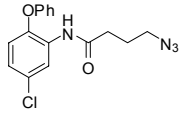
2-F12-4C



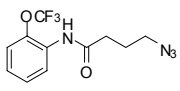
2-G1-4C



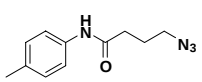
2-G2-4C



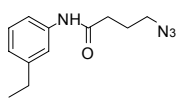
2-G3-4C



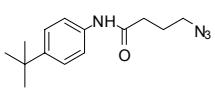
2-G4-4C



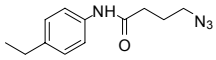
2-G5-4C



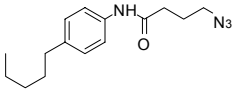
2-G6-4C



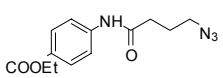
2-G7-4C



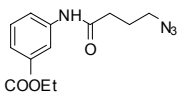
2-G8-4C



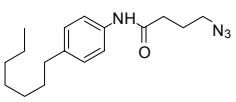
2-G9-4C



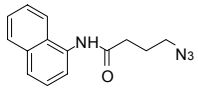
2-G10-4C



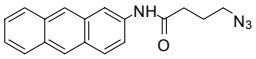
2-G11-4C



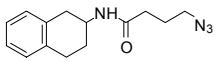
2-G12-4



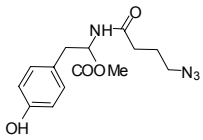
2-H1-4C



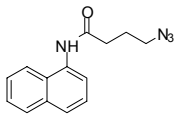
2-H2-4C



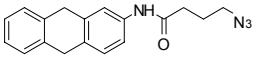
2-H3-4C



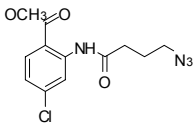
2-H4-4C



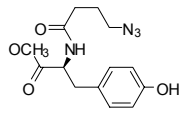
2-H5-4C



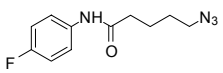
2-H7-4C



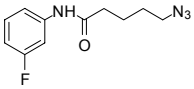
2-H8-4C



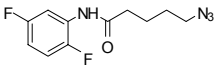
2-A1-5C



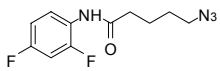
2-A2-5C



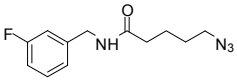
2-A3-5C



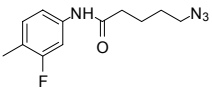
2-A3-5C



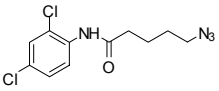
2-A5-5C



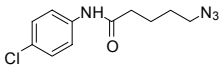
2-A5-5C



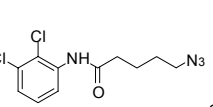
2-A7-5C



2-A8-5C

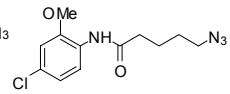


2-A9-5C



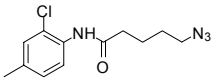
2-A9-5C

2-A11-5C



2-A12-5C

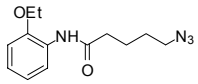
2-B1-5C



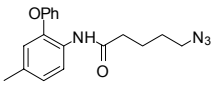
2-B2-5C

2-B3-5C

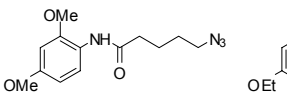
2-B4-5C



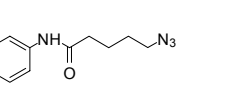
2-B5-5C



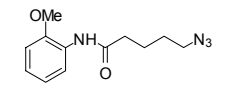
2-B6-5C



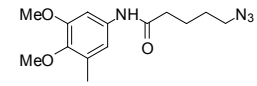
2-B7-5C



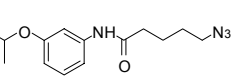
2-B8-5C



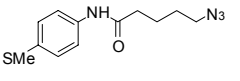
2-B9-5C



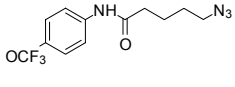
2-B10-5C



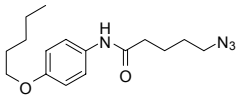
2-B11-5C



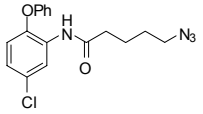
2-B12-5C



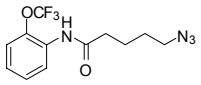
2-C1-5C



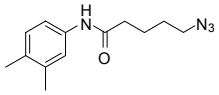
2-C2-5C



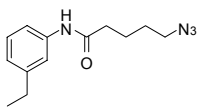
2-C3-5C



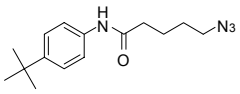
2-C4-5C



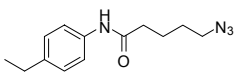
2-C5-5C



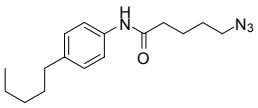
2-C6-5C



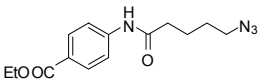
2-C7-5C



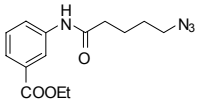
2-C8-5C



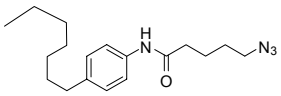
2-C9-5C



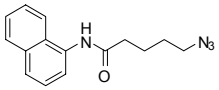
2-C10-5C



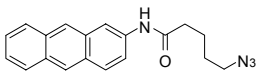
2-C11-5C



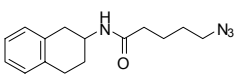
2-C12-5C



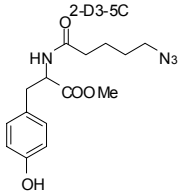
2-D1-5C



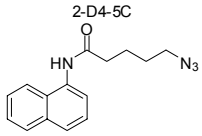
2-D2-5C



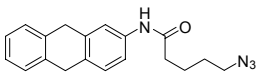
2-D3-5C



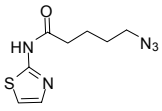
2-D4-5C



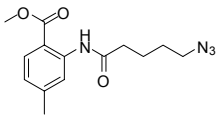
2-D5-5C



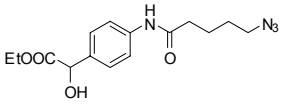
2-D6-5C



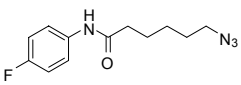
2-D7-5C



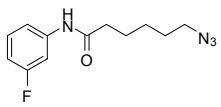
2-D8-5C



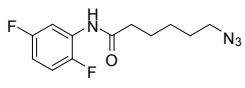
2-E1-6C



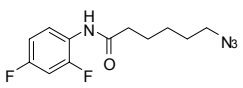
2-E2-6C



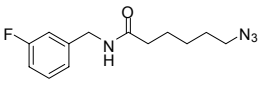
2-E3-6C



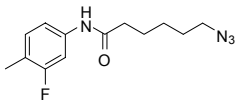
2-E4-6C



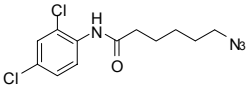
2-E5-6C



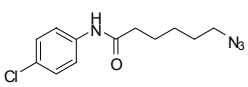
2-E6-6C



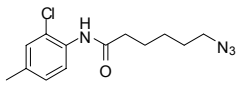
2-E7-6C



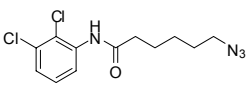
2-E8-6C



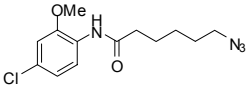
2-E9-6C



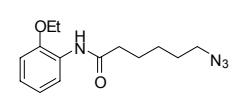
2-E10-6C



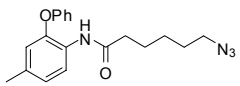
2-E11-6C



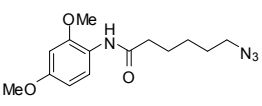
2-E12-6C



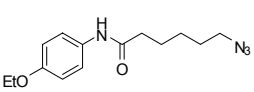
2-F1-6C



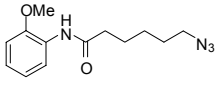
2-F2-6C



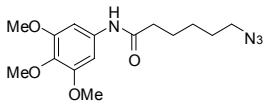
2-F3-6C



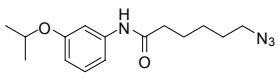
2-F4-6C



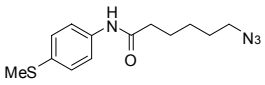
2-F5-6C



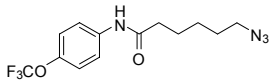
2-F6-6C



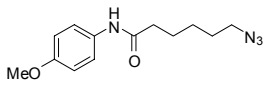
2-F7-6C



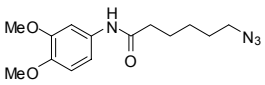
2-F8-6C



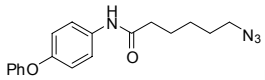
2-F9-6C



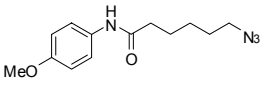
2-F10-6C



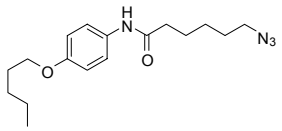
2-F11-6C



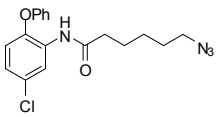
2-F12-6C



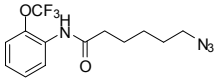
2-G1-6C



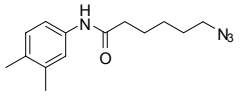
2-G2-6C



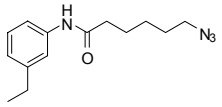
2-G3-6C



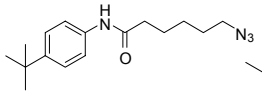
2-G4-6C



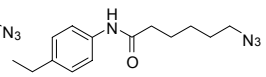
2-G5-6C



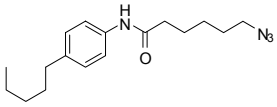
2-G6-6C



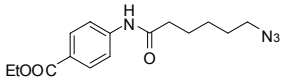
2-G7-6C



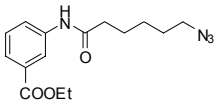
2-G8-6C



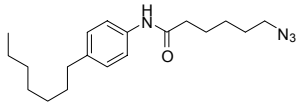
2-G9-6C



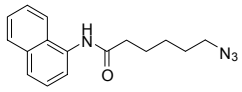
2-G10-6C



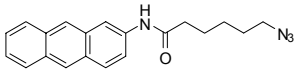
2-G11-6C



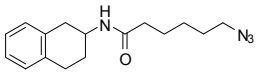
2-G12-6C



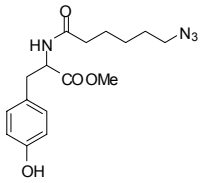
2-H1-6C



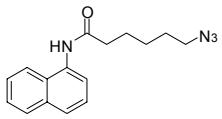
2-H2-6C



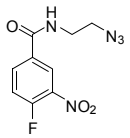
2-H3-6CH



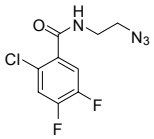
2-H4-6C



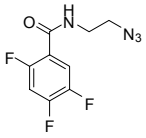
2-E1-3C



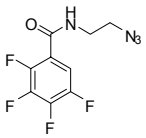
2-E2-3C



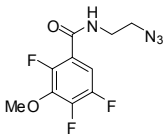
2-E3-3C



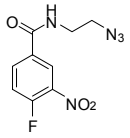
2-E4-3C



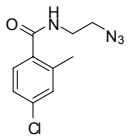
2-E5-3C



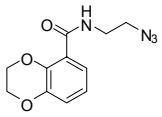
2-E6-3C



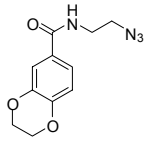
2-E7-3C



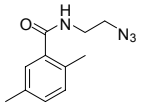
2-E8-3C



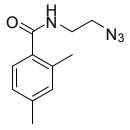
2-E9-3C



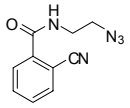
2-E10-3C



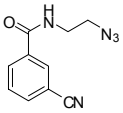
2-E11-3C



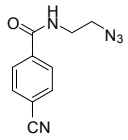
2-E12-3C



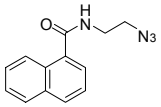
2-F1-3C



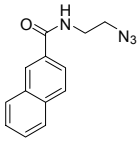
2-F2-3C



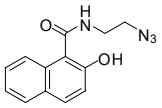
2-F3-3C



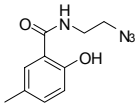
2-F4-3C



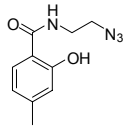
2-F5-3C



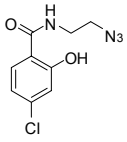
2-F6-3C



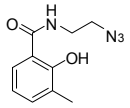
2-F7-3C



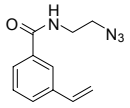
2-F8-3C



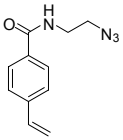
2-F9-3C



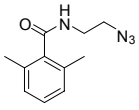
2-F10-3C



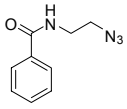
2-F11-3C



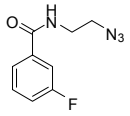
2-F12-3C



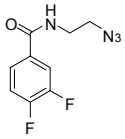
2-G1-3C



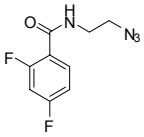
2-G2-3C



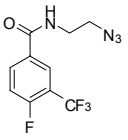
2-G3-3C



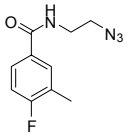
2-G4-3C



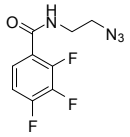
2-G5-3C



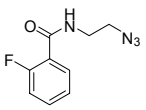
2-G6-3C



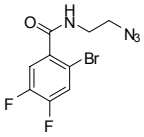
2-G7-3C



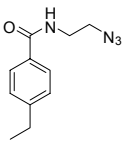
2-G8-3C



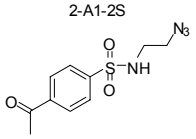
2-G9-3C



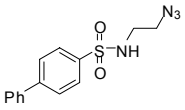
2-G10-3C



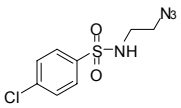
2-A1-2S



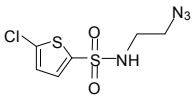
2-A2-2S



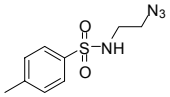
2-A3-2S



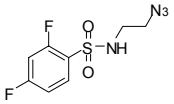
2-A4-2S



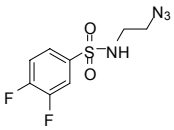
2-A5-2S



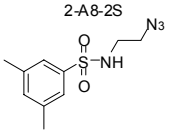
2-A6-2S



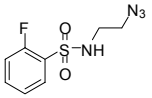
2-A7-2S



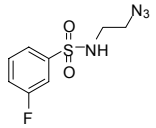
2-A8-2S



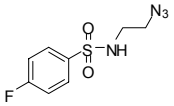
2-A9-2S



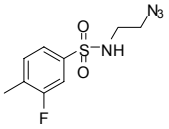
2-A10-2S



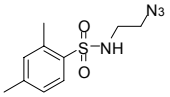
2-A11-2S



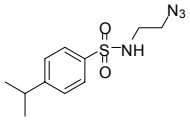
2-A12-2S



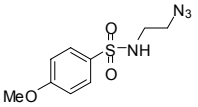
2-B1-2S



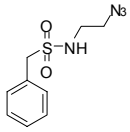
2-B2-2S



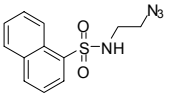
2-B3-2S



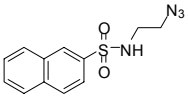
2-B4-2S



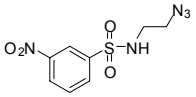
2-B5-2S



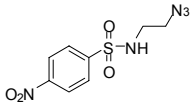
2-B6-2S



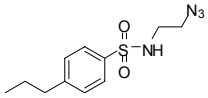
2-B7-2S



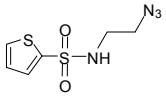
2-B8-2S



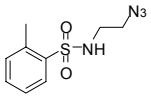
2-B9-2S



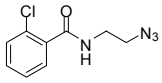
2-B10-2S



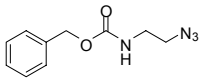
2-B11-2S



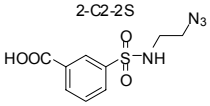
2-B12-2S



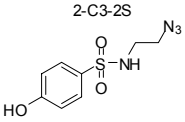
2-C1-2S



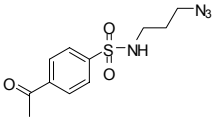
2-C2-2S



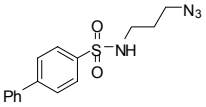
2-C3-2S



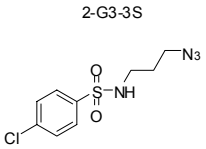
2-G1-3S



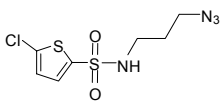
2-G2-3S



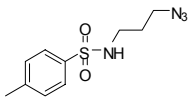
2-G3-3S



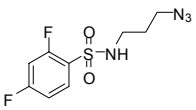
2-G4-3S



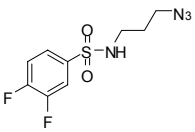
2-G5-3S



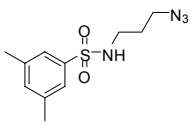
2-G6-3S



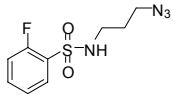
2-G7-3S



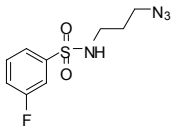
2-G8-3S



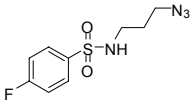
2-G9-3S



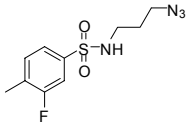
2-G10-3S



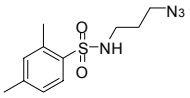
2-G11-3S



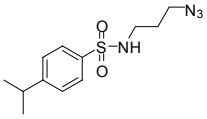
2-G12-3S



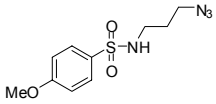
2-H1-3S



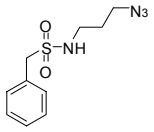
2-H2-3S



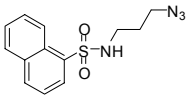
2-H3-3S



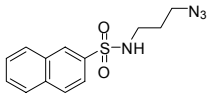
2-H4-3S



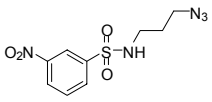
2H5-3S



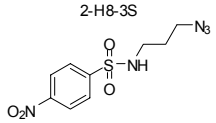
2-H6-3S



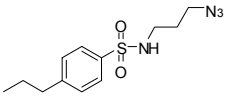
2-H7-3S



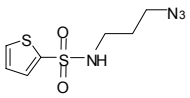
2-H8-3S



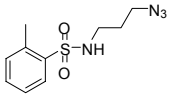
2-H9-3S



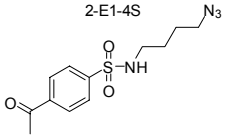
2-H10-3S



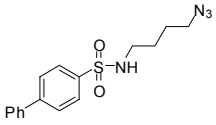
2-H11-3S



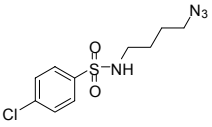
2-E1-4S



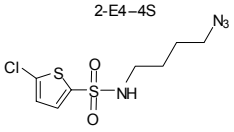
2-E2-4S



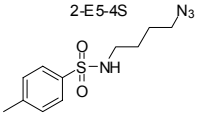
2-E3-4S



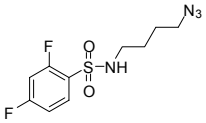
2-E4–4S



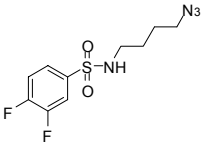
2-E5-4S



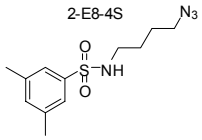
2-E6-4S



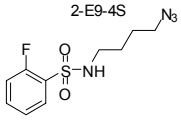
2-E7-4S



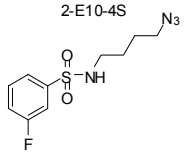
2-E8-4S



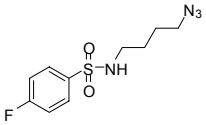
2-E9-4S



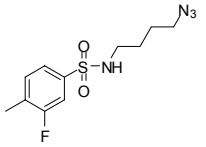
2-E10-4S



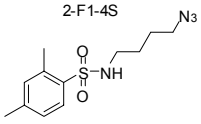
2-E11-4S



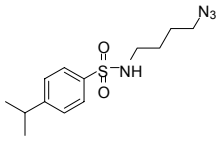
2-E12-4S



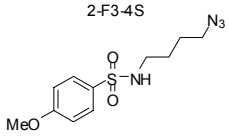
2-F1-4S



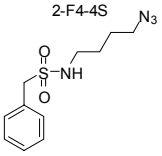
2-F2-4S



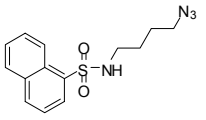
2-F3-4S



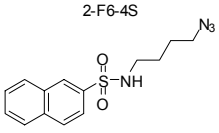
2-F4-4S



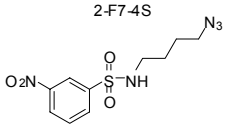
2-F5-4S



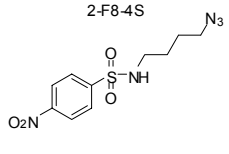
2-F6-4S



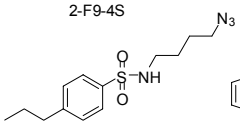
2-F7-4S



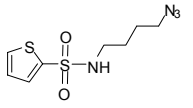
2-F8-4S



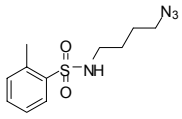
2-F9-4S



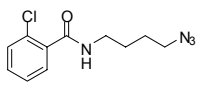
2-F10-4S



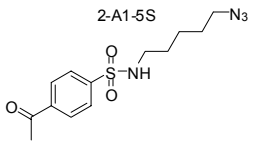
2-F11-4S



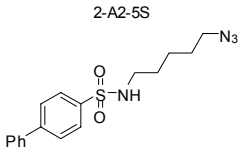
2-F12-4S



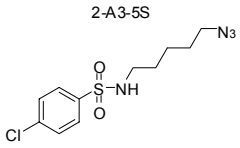
2-A1-5S



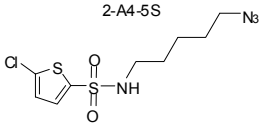
2-A2-5S



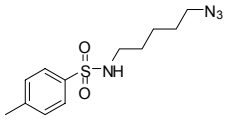
2-A3-5S



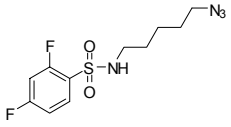
2-A4-5S



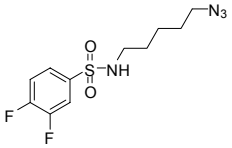
2-A5-5S



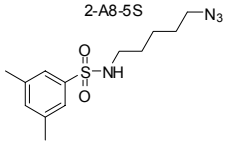
2-A6-5S



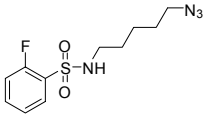
2-A7-5S



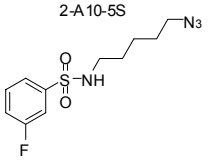
2-A8-5S



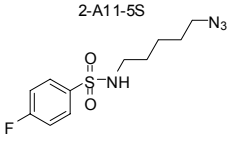
2-A9-5S



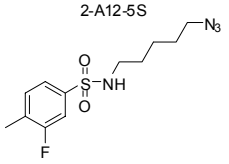
2-A10-5S



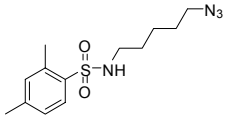
2-A11-5S



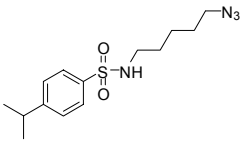
2-A12-5S



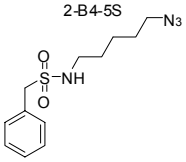
2-B1-5S



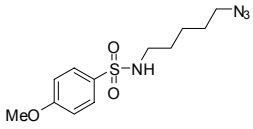
2-B2-5S



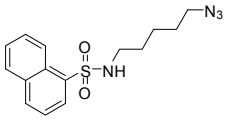
2-B4-5S



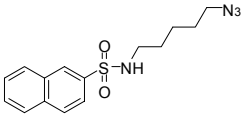
2-B3-5S



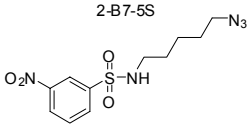
2-B5-5S



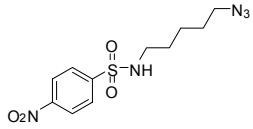
2-B6-5S



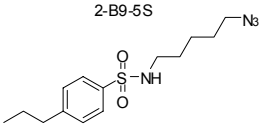
2-B7-5S



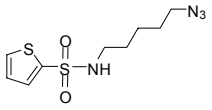
2-B8-5S



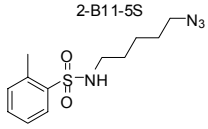
2-B9-5S



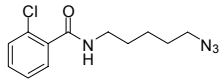
2-B10-5S



2-B11-5S



2-B12-5S



2-C1-5S

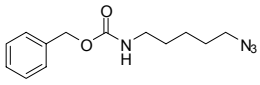


Table 10.1 Summary of traceless azide characterization

Summary of azide Characterizations					
#	ID of Azide	LCMS		^{1}H NMR	^{13}C NMR
		Est % Purity*	Obs MW		
1	2-A1-2C	>80%	N.D.	N.D	N.D
2	2-A2-2C	>95%	195.060	Y	Y
3	2-A3-2C	>60%	N.D.	N.D	N.D
4	2-A4-2C	>60%	N.D.	N.D	N.D
5	2-A5-2C	>60%	N.D.	N.D	N.D
6	2-A6-2C	>95%	N.D.	N.D	N.D
7	2-A7-2C	>95%	N.D.	N.D	N.D
8	2-A8-2C	>95%	N.D.	N.D	N.D
9	2-A9-2C	>95%	N.D.	N.D	N.D
10	2-A10-2C	>95%	N.D.	N.D	N.D
11	2-A11-2C	>95%	N.D.	N.D	N.D
12	2-A12-2C	>95%	N.D.	N.D	N.D
13	2-B1-2C	>60%	269.098	Y	Y
14	2-B2-2C	>60%	N.D.	Y	Y
15	2-B3-2C	>95%	N.D.	N.D	N.D
16	2-B4-2C	>95%	N.D.	N.D	N.D
17	2-B5-2C	>95%	267.100	Y	Y
18	2-B6-2C	>90%	N.D.	N.D	N.D
19	2-B7-2C	>95%	N.D.	N.D	N.D
20	2-B8-2C	>95%	N.D.	N.D	N.D
21	2-B9-2C	>95%	N.D.	N.D	N.D
22	2-B10-2C	>95%	N.D.	N.D	N.D
23	2-B11-2C	>95%	N.D.	N.D	N.D
24	2-B12-2C	>95%	207.080	Y	Y
25	2-C1-2C	>90%	N.D.	N.D	N.D
26	2-C2-2C	>95%	N.D.	N.D	N.D
27	2-C3-2C	>60%	N.D.	N.D	N.D
28	2-C4-2C	>95%	N.D.	Y	Y
29	2-C5-2C	>95%	205.102	N.D	N.D
30	2-C6-2C	>95%	N.D.	N.D	N.D
31	2-C7-2C	>95%	205.103	Y	Y
32	2-C8-2C	>95%	N.D.	N.D	N.D
33	2-C9-2C	>95%	N.D.	N.D	N.D
34	2-C10-2C	>95%	N.D.	Y	Y
35	2-C11-2C	>95%	275.179	Y	Y
36	2-C12-2C	>95%	277.099	N.D	N.D
37	2-D1-2C	>95%	N.D.	Y	Y
38	2-D2-2C	>90%	N.D.	N.D	N.D
39	2-D3-2C	>60%	N.D.	N.D	N.D
40	2-D4-2C	>60%	N.D.	N.D	N.D

41	2-A1-4C	>95%	N.D.	Y	Y
42	2-A2-4C	>95%	N.D.	N.D	N.D
43	2-A3-4C	>95%	N.D.	N.D	N.D
44	2-A4-4C	>95%	N.D.	N.D	N.D
45	2-A5-4C	>95%	N.D.	N.D	N.D
46	2-A6-4C	>95%	N.D.	Y	Y
47	2-A7-4C	>90%	N.D.	N.D	N.D
48	2-A8-4C	>95%	N.D.	Y	Y
49	2-A9-4C	>95%	N.D.	N.D	N.D
50	2-A10-4C	>50%	N.D.	N.D	N.D
51	2-A11-4C	>95%	N.D.	Y	Y
52	2-A12-4C	>95%	N.D.	N.D	N.D
53	2-B1-4C	>95%	N.D.	N.D	N.D
54	2-B2-4C	>95%	N.D.	N.D	N.D
55	2-B3-4C	>95%	N.D.	N.D	N.D
56	2-B4-4C	>95%	N.D.	Y	Y
57	2-B5-4C	>95%	N.D.	N.D	N.D
58	2-B6-4C	>95%	N.D.	N.D	N.D
59	2-B7-4C	>95%	N.D.	Y	Y
60	2-B8-4C	>95%	N.D.	Y	Y
61	2-B9-4C	>95%	N.D.	N.D	N.D
62	2-B10-4C	>95%	N.D.	N.D	N.D
63	2-B11-4C	>95%	N.D.	Y	Y
64	2-B12-4C	>95%	N.D.	Y	Y
65	2-C1-4C	>95%	N.D.	N.D	N.D
66	2-C2-4C	>95%	N.D.	N.D	N.D
67	2-C3-4C	-	N.D.	N.D	N.D
68	2-C4-4C	>95%	N.D.	N.D	N.D
69	2-C5-4C	>95%	N.D.	N.D	N.D
70	2-C6-4C	>95%	N.D.	Y	Y
71	2-C7-4C	>95%	N.D.	N.D	N.D
72	2-C8-4C	>95%	N.D.	N.D	N.D
73	2-C9-4C	>95%	N.D.	N.D	N.D
74	2-C10-4C	>95%	N.D.	N.D	N.D
75	2-C11-4C	>10 %	N.D.	N.D	N.D
76	2-C12-4C	>50%	N.D.	N.D	N.D
77	2-D1-4C	>95%	N.D.	N.D	N.D
78	2-D2-4C	>60%	N.D.	N.D	N.D
79	2-D3-4C	>30%	N.D.	N.D	N.D
80	2-D4-4C	>95%	N.D.	N.D	N.D
81	2-D5-4C	>90 %	N.D.	Y	N.D
82	2-D6-4C	>90 %	N.D.	Y	N.D
83	2-D7-4C	>90 %	N.D.	Y	N.D
84	2-A1-5C	>95%	237.111	N.D	N.D
85	2-A2-5C	>80%	237.106	N.D	N.D
86	2-A3-5C	>95%	255.102	N.D	N.D
87	2-A4-5C	>95%	251.124	N.D	N.D
88	2-A5-5C	>95%	251.124	N.D	N.D

89	2-A6-5C	>95%	251.127	Y	Y
90	2-A7-5C	>95%	287.038	N.D	N.D
91	2-A8-5C	>95%	253.079	Y	Y
92	2-A9-5C	>95%	267.097	N.D	N.D
93	2-A10-5C	>95%	N.D.	N.D	N.D
94	2-A11-5C	>95%	283.094	N.D	N.D
95	2-A12-5C	>95%	263.148	Y	Y
96	2-B1-5C	>90%	311.140	N.D	N.D
97	2-B2-5C	>95%	279.143	N.D	N.D
98	2-B3-5C	>95%	263.148	Y	Y
99	2-B4-5C	>95%	249.133	Y	Y
100	2-B5-5C	>95%	309.151	N.D	N.D
101	2-B6-5C	>95%	277.163	Y	Y
102	2-B7-5C	>95%	265.108	N.D	N.D
103	2-B8-5C	>95%	303.103	N.D	N.D
104	2-B9-5C	>95%	249.131	N.D	N.D
105	2-B10-5C	>95%	279.143	Y	Y
106	2-B11-5C	>95%	311.146	N.D	N.D
107	2-B12-5C	>95%	249.132	N.D	N.D
108	2-C1-5C	>95%	305.196	N.D	N.D
109	2-C2-5C	>95%	345.110	Y	Y
110	2-C3-5C	>95%	303.104	N.D	N.D
111	2-C4-5C	>95%	247.152	N.D	N.D
112	2-C5-5C	>95%	247.153	Y	Y
113	2-C6-5C	>95%	275.182	N.D	N.D
114	2-C7-5C	>90%	247.149	N.D	N.D
115	2-C8-5C	>95%	289.200	N.D	N.D
116	2-C9-5C	>95%	291.144	Y	Y
117	2-C10-5C	-	-	-	-
118	2-C11-5C	>95%	317.233	N.D	N.D
119	2-C12-5C	>95%	269.163	N.D	N.D
120	2-D1-5C	>95%	319.156	N.D	N.D
121	2-D2-5C	>50%	N.D.	N.D	N.D
122	2-D3-5C	>20%	N.D.	N.D	N.D
123	2-D4-5C	>50%	N.D.	N.D	N.D
124	2-D5-5C	>90%	N.D.	Y	N.D
125	2-D8-5C	>90%	N.D.	Y	N.D
126	2-D9-5C	>90%	N.D.	Y	N.D
127	2-D10-5C	>95%	N.D.	Y	N.D
128	2-A1-6C	>95%	251.124	Y	Y
129	2-A2-6C	>60 %	N.D.	N.D	N.D
130	2-A3-6C	>80%	N.D.	N.D	N.D
131	2-A4-6C	>80%	269.112	N.D	N.D
132	2-A5-6C	>85%	265.137	N.D	N.D
133	2-A6-6C	>95%	265.136	N.D	N.D
134	2-A7-6C	>80%	301.049	N.D	N.D
135	2-A8-6C	>95%	267.091	Y	Y
136	2-A9-6C	>80%	281.108	N.D	N.D

137	2-A10-6C	>95%	301.051	N.D	N.D
138	2-A11-6C	>95%	297.104	Y	Y
139	2-A12-6C	>95%	277.157	Y	Y
140	2-B1-6C	>60%	N.D.	N.D	N.D
141	2-B2-6C	>95%	293.151	N.D	N.D
142	2-B3-6C	>95%	277.156	Y	Y
143	2-B4-6C	>95%	263.142	Y	Y
144	2-B5-6C	>95%	323.160	Y	Y
145	2-B6-6C	>95%	291.174	N.D	N.D
146	2-B7-6C	>95%	279.120	N.D	N.D
147	2-B8-6C	-	-	-	-
148	2-B9-6C	-	-	-	-
149	2-B10-6C	>95%	279.120	Y	Y
150	2-B11-6C	>95%	293.153	Y	Y
151	2-B12-6C	>95%	263.142	N.D	N.D
152	2-C1-6C	>95%	319.205	N.D	N.D
153	2-C2-6C	-	-	-	-
154	2-C3-6C	-	-	-	-
155	2-C4-6C	>95%	261.163	N.D	N.D
156	2-C5-6C	>95%	261.163	N.D	N.D
157	2-C6-6C	>95%	289.195	N.D	N.D
158	2-C7-6C	>60%	N.D.	N.D	N.D
159	2-C8-6C	>90%	303.209	N.D	N.D
160	2-C9-6C	>95%	305.153	N.D	N.D
161	2-C10-6C	-	-	-	-
162	2-C11-6C	>95%	331.240	N.D	N.D
163	2-C12-6C	>50%	283.148	N.D	N.D
164	2-D1-6C	>95%	333.163	N.D	N.D
165	2-D2-6C	>15%	284.129	N.D	N.D
166	2-D3-6C	>40%	N.D.	N.D	N.D
167	2-D4-6C	>80%	297.165	N.D	N.D
168	2-E1-2C	>80%	254.068	N.D	N.D
169	2-E2-2C	>95%	260.063	N.D	N.D
170	2-E3-2C	>95%	263.056	N.D	N.D
171	2-E4-2C	>95%	275.077	N.D	N.D
172	2-E5-2C	>95%	275.077	N.D	N.D
173	2-E6-2C	>95%	N.D.	N.D	N.D
174	2-E7-2C	>60%	N.D.	N.D	N.D
175	2-E8-2C	>95%	249.098	N.D	N.D
176	2-E9-2C	>95%	249.098	N.D	N.D
177	2-E10-2C	>65%	219.124	N.D	N.D
178	2-E11-2C	>95%	219.125	N.D	N.D
179	2-E12-2C	>50%	216.084	N.D	N.D
180	2-F1-2C	>90%	N.D.	N.D	N.D
181	2-F2-2C	>95%	N.D.	N.D	N.D
182	2-F3-2C	>95%	241.108	N.D	N.D
183	2-F4-2C	>95%	241.108	N.D	N.D

184	2-F5-2C	>50%	N.D.	N.D	N.D
185	2-F6-2C	>50%	221.102	N.D	N.D
186	2-F7-2C	>65%	221.102	N.D	N.D
187	2-F8-2C	>60%	241.049	N.D	N.D
188	2-F9-2C	>40%	221.103	N.D	N.D
189	2-F10-2C	>95%	217.106	N.D	N.D
190	2-F11-2C	>95%	217.106	N.D	N.D
191	2-F12-2C	>40%	N.D.	N.D	N.D
192	2-G1-2C	>95%	191.092	Y	Y
193	2-G2-2C	>95%	209.081	Y	Y
194	2-G3-2C	>95%	227.073	Y	Y
195	2-G4-2C	>95%	227.072	Y	Y
196	2-G5-2C	>95%	277.068	Y	Y
197	2-G6-2C	>90%	N.D.	Y	Y
198	2-G7-2C	>95%	245.060	Y	Y
199	2-G8-2C	>95%	209.081	Y	Y
200	2-G9-2C	>80%	304.983	Y	Y
201	2-G10-2C	>95%	219.123	Y	Y
202	2-H1-2C	>95%	N.D.	N.D	N.D
203	2-H1-4C	>80%	N.D.	N.D	N.D
204	2-H1-5C	>60%	N.D.	N.D	N.D
205	2-H1-6C	>60%	N.D.	N.D	N.D
206	2-I1-2C	>95%	N.D.	Y	Y
207	2-I2-2C	>95%	N.D.	Y	Y
208	2-I3-2C	>90%	N.D.	N.D	N.D
209	2-I4-2C	>95%	N.D.	N.D	N.D
210	2-I5-2C	>85%	N.D.	N.D	N.D
211	2-I6-2C	>95%	N.D.	N.D	N.D
212	2-I7-2C	>85%	N.D.	N.D	N.D
213	2-I8-2C	>90%	N.D.	N.D	N.D
214	2-I9-2C	>85%	N.D.	N.D	N.D
215	2-I10-2C	>95%	N.D.	N.D	N.D
216	2-I11-2C	>80%	N.D.	N.D	N.D
217	2-I12-2C	>80%	N.D.	N.D	N.D
218	2-J1-2C	>85%	N.D.	N.D	N.D
219	2-J2-2C	>80%	N.D.	N.D	N.D
220	2-J3-2C	>95%	N.D.	N.D	N.D
221	2-J4-2C	-	N.D.	N.D	N.D
222	2-J5-2C	>95%	N.D.	N.D	N.D
223	2-J6-2C	>95%	N.D.	Y	Y
224	2-J7-2C	>95%	N.D.	Y	Y
225	2-J8-2C	>85%	N.D.	N.D	N.D
226	2-J9-2C	>95%	N.D.	N.D	N.D
227	2-J10-2C	>95%	N.D.	N.D	N.D
228	2-J11-2C	>70%	N.D.	N.D	N.D
229	2-J12-2C	>90%	N.D.	N.D	N.D
230	2-K1-2C	>90%	N.D.	N.D	N.D
231	2-I1-3C	>95%	N.D.	Y	Y

232	2-I2-3C	>95%	N.D.	N.D	N.D
233	2-I3-3C	>95%	N.D.	N.D	N.D
234	2-I4-3C	>95%	N.D.	Y	Y
235	2-I5-3C	>90%	N.D.	N.D	N.D
236	2-I6-3C	N.D	N.D.	N.D	N.D
237	2-I7-3C	>95%	N.D.	Y	Y
238	2-I8-3C	>85%	N.D.	N.D	N.D
239	2-I9-3C	>95%	N.D.	N.D	N.D
240	2-I10-3C	>95%	N.D.	N.D	N.D
241	2-I11-3C	N.D	N.D.	N.D	N.D
242	2-I12-3C	95%	N.D.	N.D	N.D
243	2-J1-3C	>80%	N.D.	N.D	N.D
244	2-J2-3C	>95%	N.D.	N.D	N.D
245	2-J3-3C	>95%	N.D.	N.D	N.D
246	2-J4-3C	N.D	N.D.	N.D	N.D
247	2-J5-3C	>95%	N.D.	N.D	N.D
248	2-J6-3C	>95%	N.D.	Y	Y
249	2-J7-3C	>95%	N.D.	N.D	N.D
250	2-J8-3C	-	N.D.	N.D	N.D
251	2-J9-3C	>95%	N.D.	N.D	N.D
252	2-J10-3C	>95%	N.D.	N.D	N.D
253	2-J11-3C	>90%	N.D.	N.D	N.D
254	2-I1-4C	>95%	N.D.	Y	Y
255	2-I2-4C	>95%	N.D.	Y	Y
256	2-I3-4C	>90%	N.D.	N.D	N.D
257	2-I4-4C	>95%	N.D.	N.D	N.D
258	2-I5-4C	>85%	N.D.	N.D	N.D
259	2-I6-4C	>85%	N.D.	N.D	N.D
260	2-I7-4C	>85%	N.D.	N.D	N.D
261	2-I8-4C	>75%	N.D.	N.D	N.D
262	2-I9-4C	>80%	N.D.	N.D	N.D
263	2-I10-4C	>90%	N.D.	N.D	N.D
264	2-I11-4C	>70%	N.D.	N.D	N.D
265	2-I12-4C	>90%	N.D.	N.D	N.D
266	2-J1-4C	>70%	N.D.	N.D	N.D
267	2-J2-4C	>80%	N.D.	N.D	N.D
268	2-J3-4C	>95%	N.D.	N.D	N.D
269	2-J4-4C	>95%	N.D.	N.D	N.D
270	2-J5-4C	>95%	N.D.	Y	Y
271	2-J6-4C	>95%	N.D.	Y	Y
272	2-J7-4C	>95%	N.D.	N.D	N.D
273	2-J8-4C	>90%	N.D.	N.D	N.D
274	2-J9-4C	>95%	N.D.	N.D	N.D
275	2-J10-4C	>70%	N.D.	N.D	N.D
276	2-J11-4C	>80%	N.D.	N.D	N.D
277	2-I1-5C	>95%	N.D.	Y	Y
278	2-I2-5C	>95%	N.D.	Y	Y
279	2-I3-5C	>95%	N.D.	N.D	N.D

280	2-I4-5C	>95%	N.D.	N.D	N.D
281	2-I5-5C	>80%	N.D.	N.D	N.D
282	2-I6-5C	>85%	N.D.	N.D	N.D
283	2-I7-5C	>85%	N.D.	N.D	N.D
284	2-I8-5C	>60%	N.D.	N.D	N.D
285	2-I9-5C	>90%	N.D.	N.D	N.D
286	2-I10-5C	>90%	N.D.	N.D	N.D
287	2-I11-5C	>75%	N.D.	N.D	N.D
288	2-I12-5C	>90%	N.D.	N.D	N.D
289	2-J1-5C	>75%	N.D.	N.D	N.D
290	2-J2-5C	>90%	N.D.	N.D	N.D
291	2-J3-5C	>95%	N.D.	N.D	N.D
292	2-J4-5C	-	N.D.	N.D	N.D
293	2-J5-5C	>95%	N.D.	N.D	N.D
294	2-J6-5C	>95%	N.D.	Y	Y
295	2-J7-5C	>95%	N.D.	Y	Y
296	2-J8-5C	>95%	N.D.	N.D	N.D
297	2-J9-5C	>95%	N.D.	N.D	N.D
298	2-J10-5C	>85%	N.D.	N.D	N.D
299	2-J11-5C	>70%	N.D.	N.D	N.D
300	2-I1-6C	>95%	N.D.	Y	Y
301	2-I2-6C	>95%	N.D.	Y	Y
302	2-I3-6C	>95%	N.D.	N.D	N.D
303	2-I4-6C	>95%	N.D.	N.D	N.D
304	2-I5-6C	>95%	N.D.	N.D	N.D
305	2-I6-6C	>90%	N.D.	N.D	N.D
306	2-I7-6C	>95%	N.D.	N.D	N.D
307	2-I8-6C	>90%	N.D.	N.D	N.D
308	2-I9-6C	>90%	N.D.	N.D	N.D
309	2-I10-6C	>95%	N.D.	N.D	N.D
310	2-I11-6C	>60%	N.D.	N.D	N.D
311	2-I12-6C	>95%	N.D.	N.D	N.D
312	2-J1-6C	>75%	N.D.	N.D	N.D
313	2-J2-6C	>90%	N.D.	N.D	N.D
314	2-J3-6C	>95%	N.D.	N.D	N.D
315	2-J4-6C	-	N.D.	N.D	N.D
316	2-J5-6C	>95%	N.D.	N.D	N.D
317	2-J6-6C	>95%	N.D.	Y	Y
318	2-J7-6C	>95%	N.D.	Y	Y
319	2-J8-6C	>95%	N.D.	N.D	N.D
320	2-J9-6C	>95%	N.D.	N.D	N.D
321	2-J10-6C	>95%	N.D.	N.D	N.D
322	2-J11-6C	>60%	N.D.	N.D	N.D
323	2-L1-2C	>80%	N.D.	N.D	N.D
324	2-L1-5C	>85 %	N.D.	N.D	N.D
325	2-L1-6C	>90%	N.D.	N.D	N.D

	Color Code		**Building Blocks**		
			Amine		
			Solution phase synthesis (amine/sulfonyl chloride)		
			Acid		
			Acid chloride		
			Sulfonyl chloride		
			Chlorformate		
			"-" = No characterizations (LCMS/NMR) were done "N.D." = Although the product is correct, the molecular mass did not show up in the mass spectrometry (due to low ionization which is		
			A - D IS AMINE BUILDING BLOCKS		
			E - G IS ACID BUILDING BLOCKS		
			H IS ACID CHLORIDE BUILDING BLOCKS		
			I - K SULFONYL CHLORIDE BUILDING BLOCKS		
			L IS CHLOROFORMATE BUILDING BLOCKS		

Table 10.2 Observed molecular weight [MS (TOF)] of the representative triazole products of the 3250-member library

Summary of Characterization of Triazole products					
#	Triazole ID	Obs MS	#	Triazole ID	Obs MS
1	**2-A-2-A1-5C**	480.148	74	**2-G-2-A7-5C**	544.111
2	**2-A-2-A3-5C**	498.136	75	**2-G-2-A9-5C**	524.167
3	**2-A-2-A4-5C**	498.136	76	**2-G-2-A10-5C**	544.115
4	**2-A-2-A5-5C**	494.164	77	**2-G-2-B2-5C**	536.208
5	**2-A-2-A8-5C**	496.114	78	**2-G-2-B7-5C**	522.172
6	**2-A-2-A9-5C**	510.135	79	**2-G-2-B8-5C**	560.171
7	**2-A-2-A10-5C**	530.08	80	**2-G-2-B1-2S**	516.131
8	**2-A-2-A11-5C**	526.127	81	**2-G-2-B1-5S**	558.179
9	**2-A-2-A12-5C**	506.177	82	**2-G-2-E1-3C**	511.129
10	**2-A-2-B1-2C**	512.131	83	**2-G-2-B6-2C**	520.103
11	**2-A-2-B2-2C**	480.129	84	**2-G-2-A1-5S**	568.172
12	**2-A-2-B3-2C**	464.133	85	**2-G-2-A2-2S**	560.148
13	**2-A-2-B4-2C**	450.12	86	**2-G-2-A2-5S**	602.195
14	**2-A-2-B5-2C**	510.137	87	**2-H-2-A2-2C**	595.927
15	**2-A-2-B6-2C**	478.153	88	**2-H-2-A2-2C**	595.921
16	**2-A-2-B7-2C**	466.094	89	**2-H-2-A4-2C**	613.919
17	**2-A-2-B8-2C**	504.089	90	**2-H-2-A5-2C**	609.948
18	**2-A-2-B9-2C**	450.118	91	**2-H-2-A9-2C**	625.914
19	**2-A-2-B10-2C**	480.134	92	**2-H-2-A11-2C**	641.908
20	**2-A-2-B11-2C**	512.133	93	**2-H-2-A12-2C**	621.966
21	**2-A-2-B12-2C**	450.12	94	**2-H-2-E2-3C**	661.887
22	**2-B-2-A1-5C**	514.125	95	**2-H-2-E3-3C**	645.923
23	**2-B-2-A2-5C**	514.125	96	**2-H-2-E4-3C**	663.923
24	**2-B-2-A3-5C**	532.118	97	**2-H-2-E6-3C**	639.943
25	**2-B-2-A4-5C**	532.117	98	**2-H-2-E7-3C**	649.958
26	**2-B-2-A5-5C**	528.144	99	**2-H-2-E12-3C**	617.922
27	**2-B-2-A6-5C**	528.145	100	**2-H-2-A1-5C**	637.984
28	**2-B-2-A8-5C**	530.096	101	**2-H-2-A2-5C**	637.978
29	**2-B-2-B1-2S**	536.048	102	**2-H-2-A3-5C**	655.967
30	**2-B-2-B2-5S**	546.089	103	**2-H-2-A4-5C**	655.968
31	**2-C-2-A1-5C**	498.156	104	**2-H-2-A5-5C**	629.989
32	**2-C-2-A2-5C**	498.158	105	**2-H-2-A4-5C**	655.968
33	**2-C-2-A3-5C**	516.144	106	**2-I-2-A1-2C**	456.101
34	**2-C-2-A4-5C**	516.144	107	**2-I-2-A2-2C**	456.104
35	**2-C-2-A5-5C**	512.17	108	**2-I-2-A4-2C**	474.101
36	**2-C-2-A9-5C**	528.148	109	**2-I-2-A5-2C**	470.127

37	**2-C-2-A10-5C**	548.088	110	**2-I-2-A6-2C**	470.125
38	**2-C-2-E1-6C**	512.172	111	**2-I-2-A8-2C**	472.082
39	**2-C-2-E2-6C**	512.172	112	**2-I-2-A9-2C**	486.098
40	**2-C-2-E3-6C**	530.164	113	**2-I-2-A11-2C**	502.098
41	**2-C-2-E4-6C**	530.162	114	**2-I-2-A12-2C**	482.154
42	**2-C-2-E5-6C**	526.191	115	**2-I-2-B1-2C**	530.15
43	**2-C-2-E6-6C**	526.152	116	**2-I-2-B2-2C**	498.148
44	**2-C-2-A9-5C**	528.148	117	**2-I-2-B3-2C**	482.145
45	**2-D-2-A1-5C**	578.234	118	**2-I-2-B4-2C**	468.129
46	**2-D-2-A4-5C**	596.2	119	**2-I-2-B5-2C**	528.159
47	**2-D-2-A6-5C**	592.212	120	**2-I-2-B6-2C**	496.163
48	**2-E-2-A1-2C**	471.086	121	**2-I-2-B7-2C**	484.104
49	**2-E-2-A2-2C**	471.097	122	**2-I-2-B8-2C**	522.101
50	**2-E-2-A5-2C**	485.114	123	**2-J-2-A1-2C**	438.122
51	**2-E-2-E1-4C**	499.129	124	**2-J-2-A2-2C**	438.125
52	**2-E-2-A1-5C**	513.143	125	**2-J-2-A4-2C**	456.119
53	**2-E-2-A2-5C**	513.139	126	**2-J-2-A5-2C**	452.139
54	**2-E-2-A3-5C**	531.135	127	**2-J-2-A6-2C**	452.145
55	**2-E-2-A4-5C**	531.13	128	**2-J-2-A8-2C**	454.098
56	**2-E-2-A5-5C**	527.155	129	**2-J-2-A9-2C**	468.112
57	**2-F-2-A6-2C**	482.133	130	**2-J-2-A11-2C**	484.153
58	**2-F-2-A8-2C**	484.088	131	**2-J-2-A12-2C**	464.163
59	**2-F-2-A9-2C**	498.107	132	**2-J-2-E4-4C**	484.153
60	**2-F-2-A11-2C**	514.109	133	**2-J-2-E5-4C**	480.176
61	**2-F-2-A12-2C**	494.163	134	**2-J-2-E6-4C**	480.178
62	**2-F-2-B3-2C**	494.158	135	**2-J-2-A1-5C**	480.153
63	**2-F-2-B4-2C**	480.147	136	**2-J-2-A2-5C**	480.153
64	**2-F-2-B5-2C**	540.164	137	**2-J-2-A3-5C**	480.163
65	**2-F-2-A1-5C**	510.178	138	**2-J-2-A4-5C**	498.15
66	**2-F-2-A2-5C**	510.173	139	**2-J-2-A5-5C**	498.153
67	**2-G-2-A1-2C**	452.139	140	**2-J-2-A6-5C**	494.183
68	**2-G-2-A2-2C**	452.141	141	**2-J-2-A11-5C**	526.146
69	**2-G-2-A1-5C**	494.179	142	**2-J-2-A12-5C**	506.2
70	**2-G-2-A2-5C**	494.177	143	**2-J-2-E1-6C**	494.176
71	**2-G-2-A3-5C**	512.168	144	**2-J-2-E2-6C**	494.176
72	**2-G-2-A4-5C**	512.169	145	**2-J-2-E3-6C**	512.163
73	**2-G-2-A7-5C**	544.111	146	**2-J2-E4-6C**	512.168

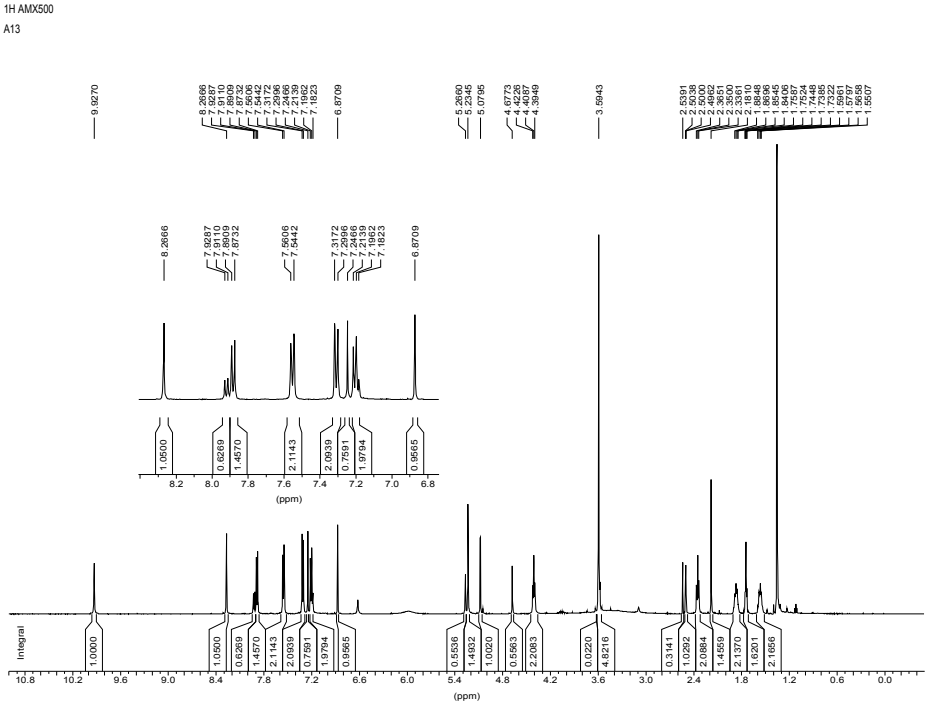
Table 10.3 Summary of characterization and IC_{50} of traceless PTP1B inhibitors

Summary Of Characterization and IC_{50} of Traceless PTP inhibitors							
#	Product ID	Acid used	Amine used	Est % Puriy	Obs MS	NMR (1H)	IC_{50} against PTP1B
1	**3-A1**	**3-A**	**3-a1**	-	-	Yes	115.3
2	**3-A2**	**3-A**	**3-a2**	-	-	Yes	167.8
3	**3-A3**	**3-A**	**3-a3**	-	-	Yes	0
4	**3-A4**	**3-A**	**3-a4**	>95 %	-	-	106.8
5	**3-A5**	**3-A**	**3-a5**	>95 %	463	-	216.4
6	**3-A6**	**3-A**	**3-a6**	>95 %	412	-	111
7	**3-A7**	**3-A**	**3-a7**	>95 %	520	-	77.3
8	**3-A8**	**3-A**	**3-a8**	-	-	-	177.1
9	**3-A9**	**3-A**	**3-a9**	>95 %	-	-	125.9
10	**3-A10**	**3-A**	**3-a10**	>95 %	438	-	395.3
11	**3-A11**	**3-A**	**3-a11**	>95 %	482	-	66.6
12	**3-A12**	**3-A**	**3-a12**	-	-	Yes	
13	**3-A13**	**3-A**	**3-a13**	>95 %	425	-	120.9
14	**3-A14**	**3-A**	**3-a14**	>95 %	484	-	139.4
15	**3-A15**	**3-A**	**3-a15**	-	-	-	60.8
16	**3-A16**	**3-A**	**3-a16**	>95 %	478	Yes	145.9
17	**3-A17**	**3-A**	**3-a17**	>95 %	440	-	115.9
18	**3-A18**	**3-A**	**3-a18**	>60 %	-	-	>400
19	**3-A19**	**3-A**	**3-a19**	>95 %	439	-	182.9
20	**3-A20**	**3-A**	**3-a20**	>95 %	424	-	216.5
21	**3-A21**	**3-A**	**3-a21**	-	-	-	27.6
22	**3-A22**	**3-A**	**3-a22**	-	-	-	148.7
23	**3-A23**	**3-A**	**3-a23**	-	-	Yes	0
24	**3-A24**	**3-A**	**3-a24**	>95 %	472	-	99.4
25	**3-A25**	**3-A**	**3-a25**	>50 %	-	-	192.6
26	**3-A26**	**3-A**	**3-a26**	>80 %	494	-	8.8
27	**3-A27**	**3-A**	**3-a27**	>95 %	467	Yes	68
28	**3-A28**	**3-A**	**3-a28**	>95 %	466	-	157.9
29	**3-A29**	**3-A**	**3-a29**	>95 %	424	-	75.5
30	**3-A30**	**3-A**	**3-a30**	>95 %	464	Yes	75
31	**3-A31**	**3-A**	**3-a31**	>95 %	423	Yes	147.2
32	**3-A32**	**3-A**	**3-a32**	>90 %	466	-	150.7
33	**3-A33**	**3-A**	**3-a33**	>95 %	512	-	0
34	**3-A34**	**3-A**	**3-a34**	-	-	-	20.6
35	**3-A35**	**3-A**	**3-a35**	-	-	-	34.2

36	**3-B1**	**3-B**	**3-a36**	-	-	-	>400
37	**3-B2**	**3-B**	**3-a37**	>70 %	454	-	
38	**3-B3**	**3-B**	**3-a38**	-	-	-	>400
39	**3-B4**	**3-B**	**3-a39**	-	-	Yes	>400
40	**3-B5**	**3-B**	**3-a40**	-	-	-	236.3
41	**3-B6**	**3-B**	**3-a41**	>60 %	440	-	>400
42	**3-B7**	**3-B**	**3-a42**	>90 %	-	-	36.5
43	**3-B8**	**3-B**	**3-a43**	-	-	-	58.7
44	**3-B9**	**3-B**	**3-a44**	-	-	Yes	>400
45	**3-B10**	**3-B**	**3-a45**	>95 %	467	Yes	>400
46	**3-B11**	**3-B**	**3-a46**	>95 %	509	-	0
47	**3-B12**	**3-B**	**3-a47**	-	-	-	183.2
48	**3-B13**	**3-B**	**3-a48**	>95 %	452	-	204.2
49	**3-B14**	**3-B**	**3-a49**	>95 %	512	-	256.8
50	**3-B15**	**3-B**	**3-a50**	>95 %	481	-	>400
51	**3-B16**	**3-B**	**3-a51**	>95 %	529	-	174.1
52	**3-B17**	**3-B**	**3-a52**	>95 %	469	Yes	>400
53	**3-B18**	**3-B**	**3-a53**	-	-	-	>400
54	**3-B19**	**3-B**	**3-a54**	>95 %	467	-	304.7
55	**3-B20**	**3-B**	**3-a55**	-	-	-	149.7
56	**3-B21**	**3-B**	**3-a56**	>95 %	524	Yes	105.8
57	**3-B22**	**3-B**	**3-a57**	-	-	-	275.2
58	**3-B23**	**3-B**	**3-a58**	>95 %	487	Yes	117.2
59	**3-B24**	**3-B**	**3-a59**	>95 %	500	-	365.9
60	**3-B25**	**3-B**	**3-a60**	-	-	-	20.5
61	**3-B26**	**3-B**	**3-a61**	-	-	-	197.7
62	**3-B27**	**3-B**	**3-a62**	-	-	-	378.8
63	**3-B28**	**3-B**	**3-a63**	-	-	-	244.5
64	**3-B29**	**3-B**	**3-a64**	>95 %	451	-	>400
65	**3-B30**	**3-B**	**3-a65**	-	-	-	>400
66	**3-B31**	**3-B**	**3-a66**	>95 %	451	-	>400
67	**3-B32**	**3-B**	**3-a67**	-	-	-	379.3
68	**3-B33**	**3-B**	**3-a68**	-	-	-	>400
69	**3-B34**	**3-B**	**3-a69**	-	-	-	>400
70	**3-B35**	**3-B**	**3-a70**	>95 %	493	-	209.8

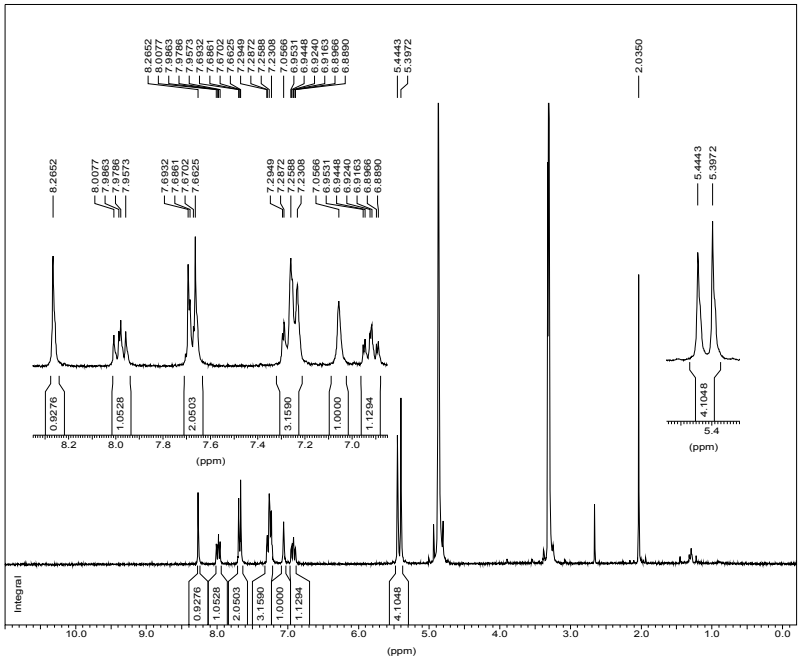
10.2 NMR Spectra for selected compounds

Compound 2-A-13



Compound 2-B8-2C-2-I

1H normal range AC300, B08-MeOD

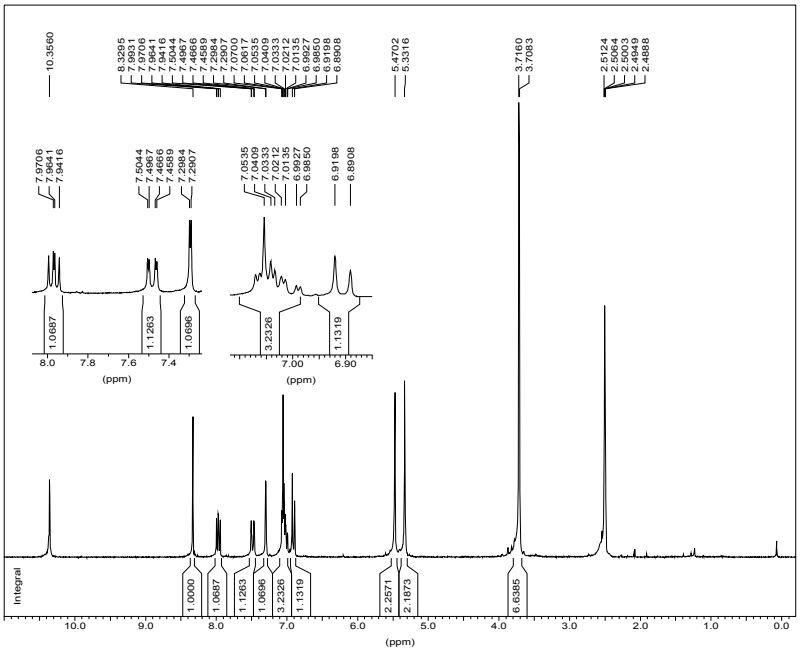


*** Current Data Parameters ***

NAME	:	ag06kka
EXPNO	:	1
PROCNO	:	1
*** Acquisition Parameters ***		
DATE_t	:	03:58:43
DATE_d	:	Aug 07 2007
NS	:	32
O1	:	1853.43 Hz
O2	:	1853.43 Hz
O3	:	1853.43 Hz
SFO1	:	300.1318534 MHz
SFO2	:	300.1318534 MHz
SFO3	:	300.1318534 MHz
*** 1D NMR Plot Parameters ***		
SR	:	4.70 Hz

Compound 2-B10-2C-I

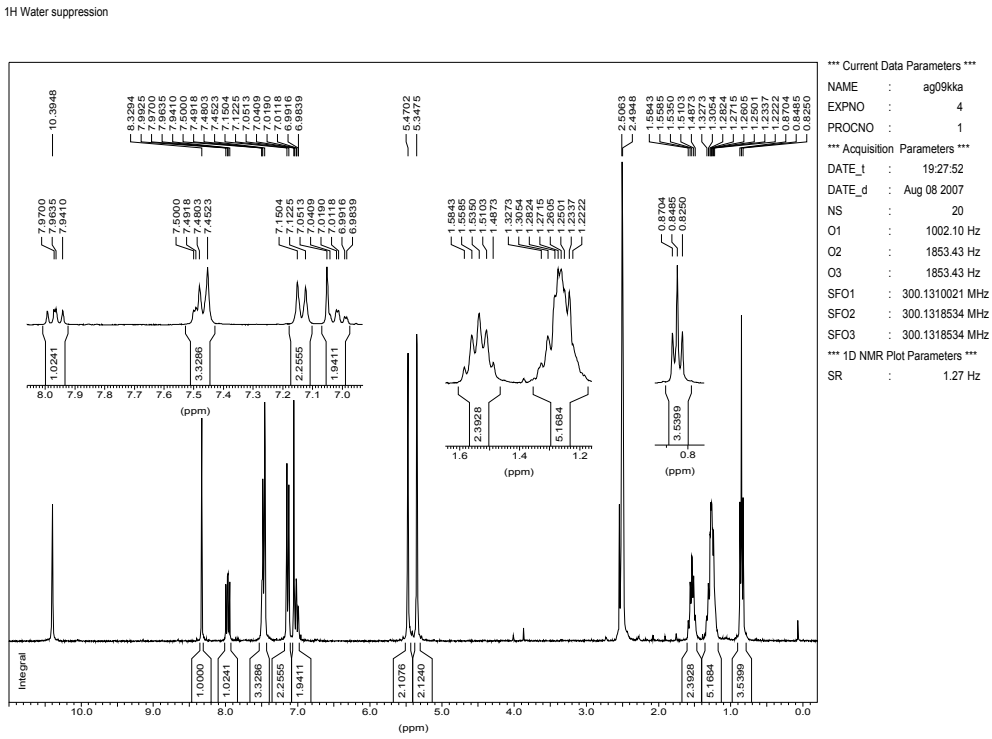
1H Water suppression, B10



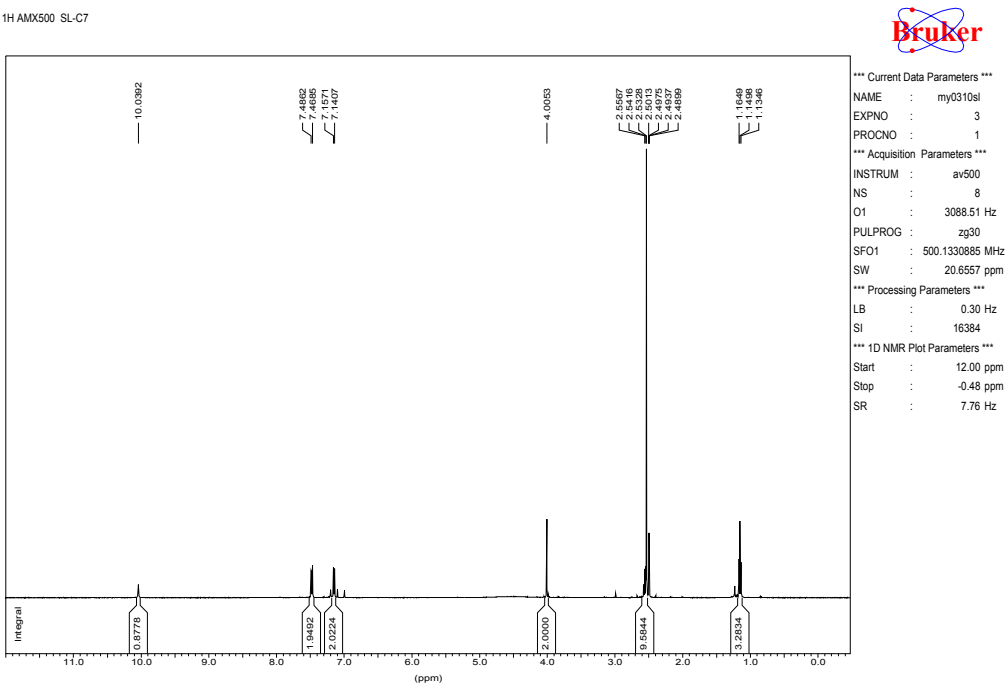
*** Current Data Parameters ***

NAME	:	ag09kka
EXPNO	:	12
PROCNO	:	1
*** Acquisition Parameters ***		
DATE_t	:	19:48:25
DATE_d	:	Aug 08 2007
NS	:	20
O1	:	1002.80 Hz
O2	:	1853.43 Hz
O3	:	1853.43 Hz
SFO1	:	300.1310028 MHz
SFO2	:	300.1318534 MHz
SFO3	:	300.1318534 MHz
*** 1D NMR Plot Parameters ***		
SR	:	0.80 Hz

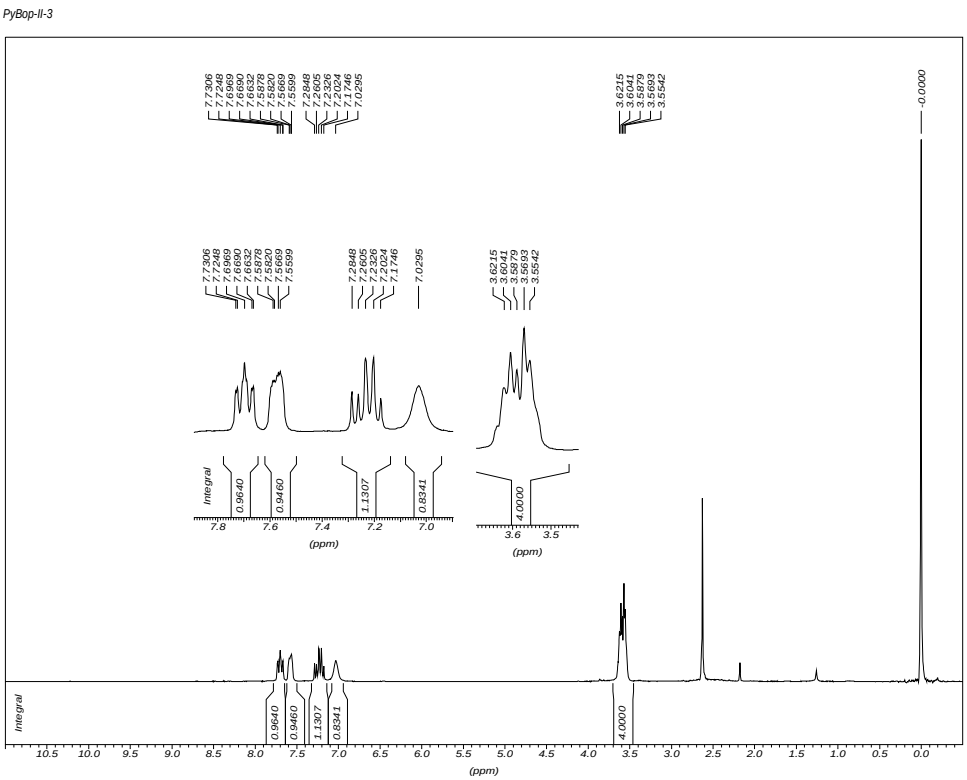
Compound 2-C8-2C-2-I



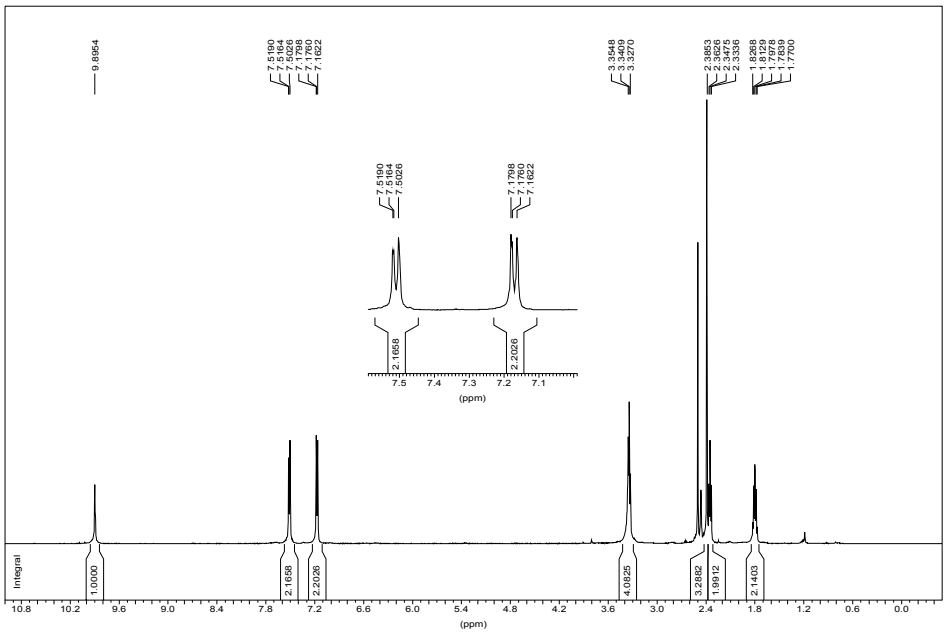
Compound 2-C7-2C



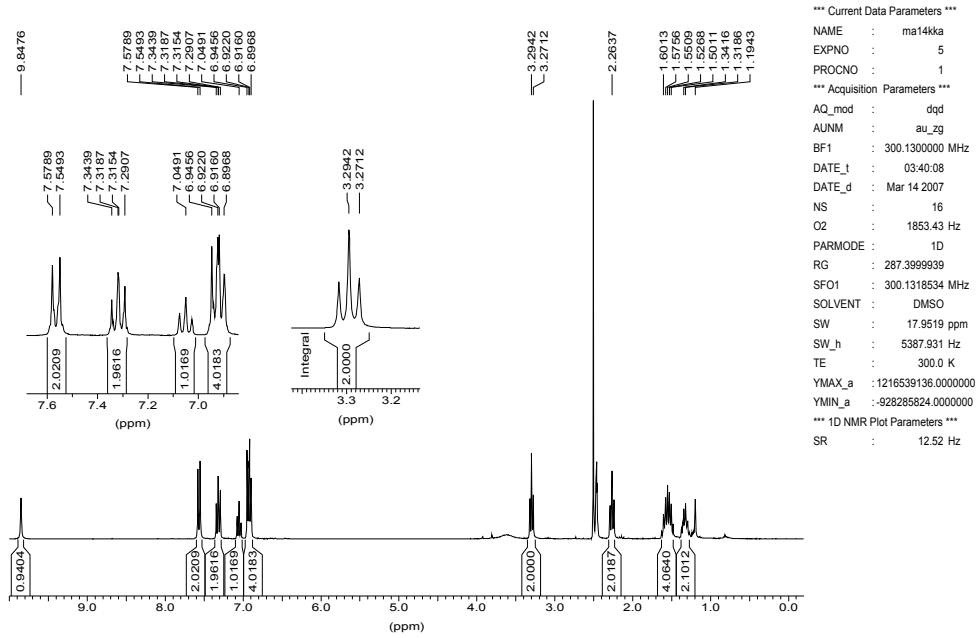
Compound 2-G3-3C



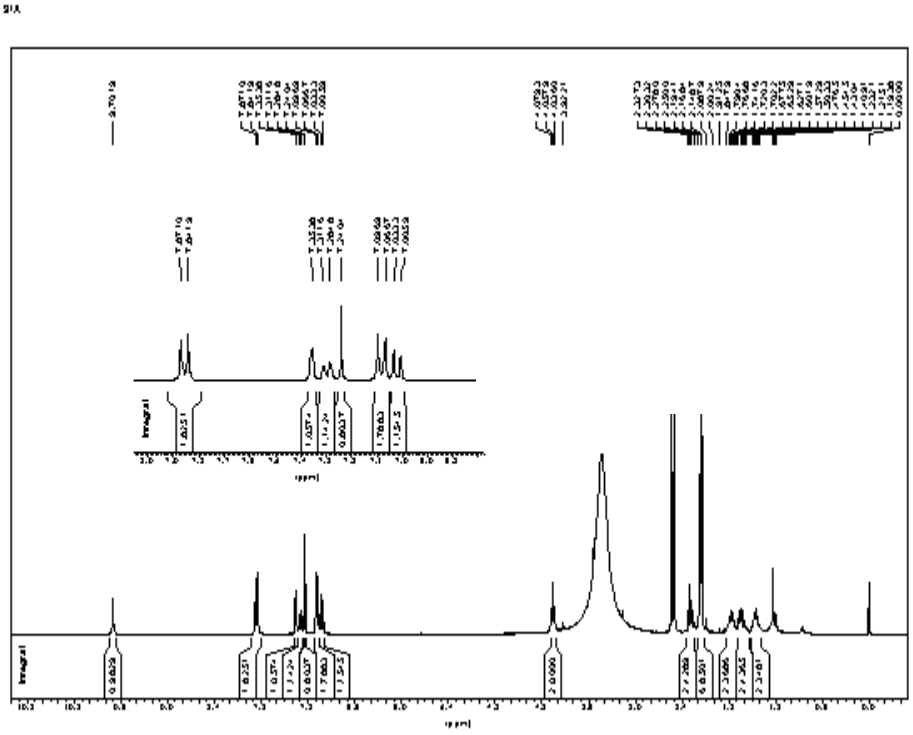
Compound 2B-2C



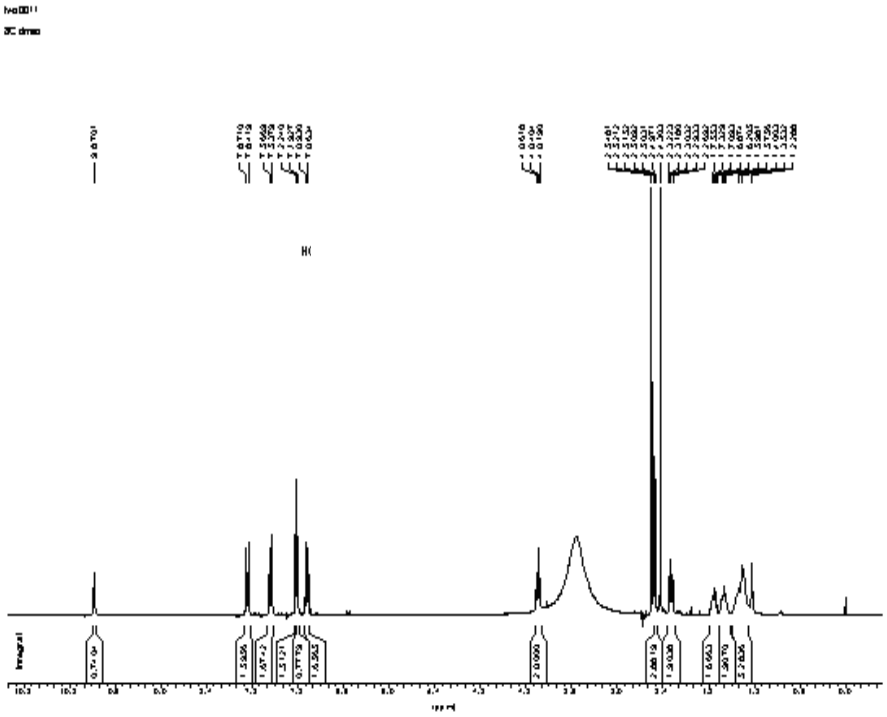
Compound 2-F11-6C



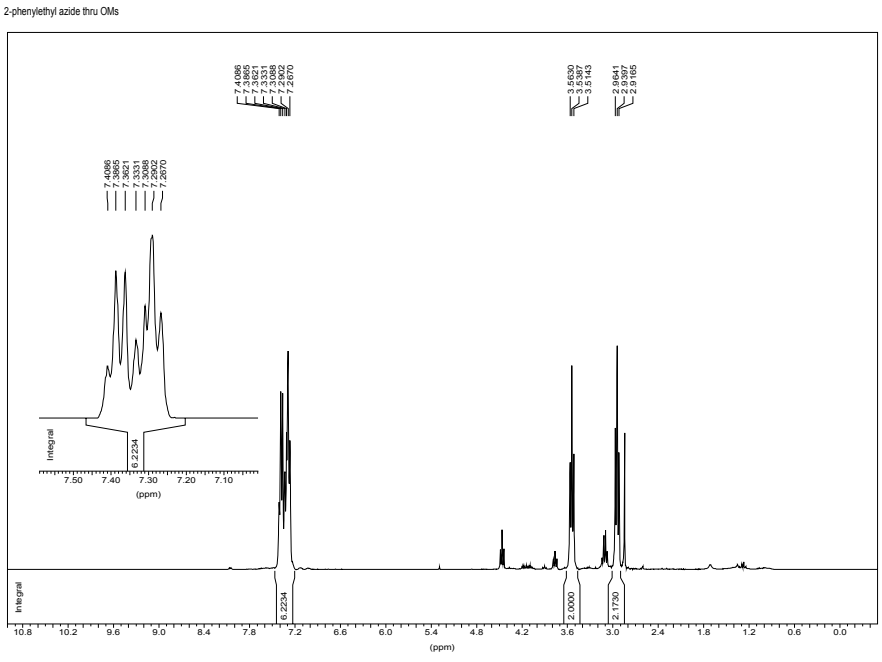
Compound 3-31A



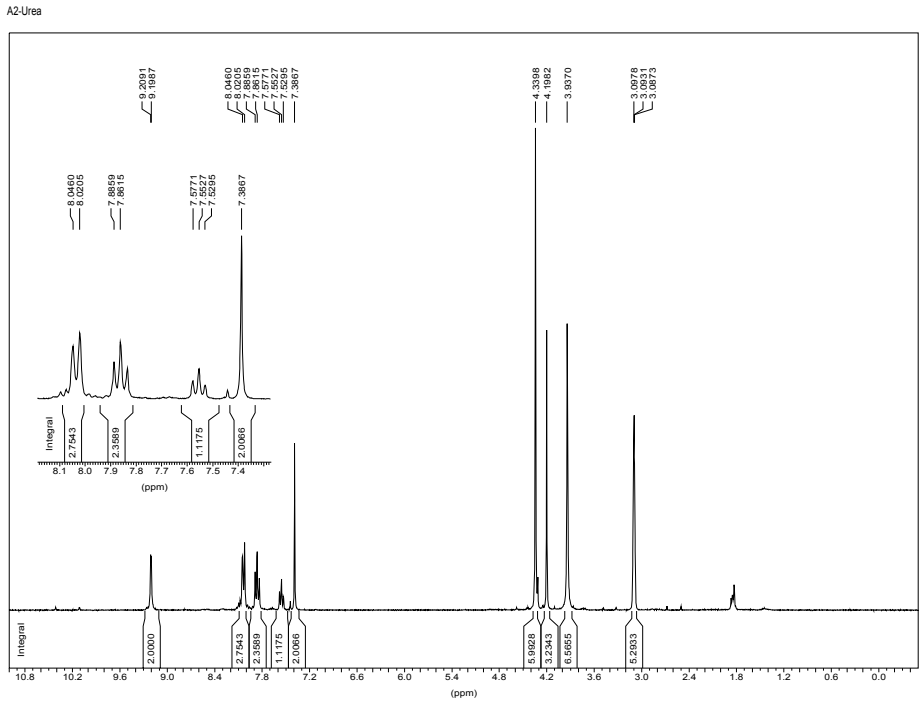
Compound 3-16B



Compound 4-5

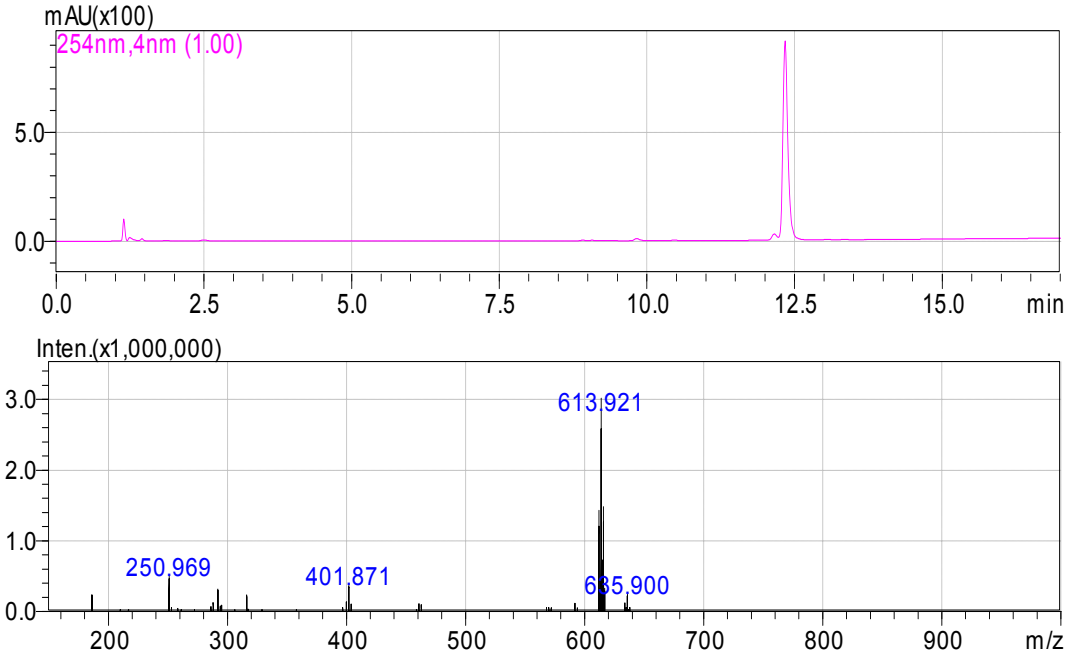


Compound 4-22

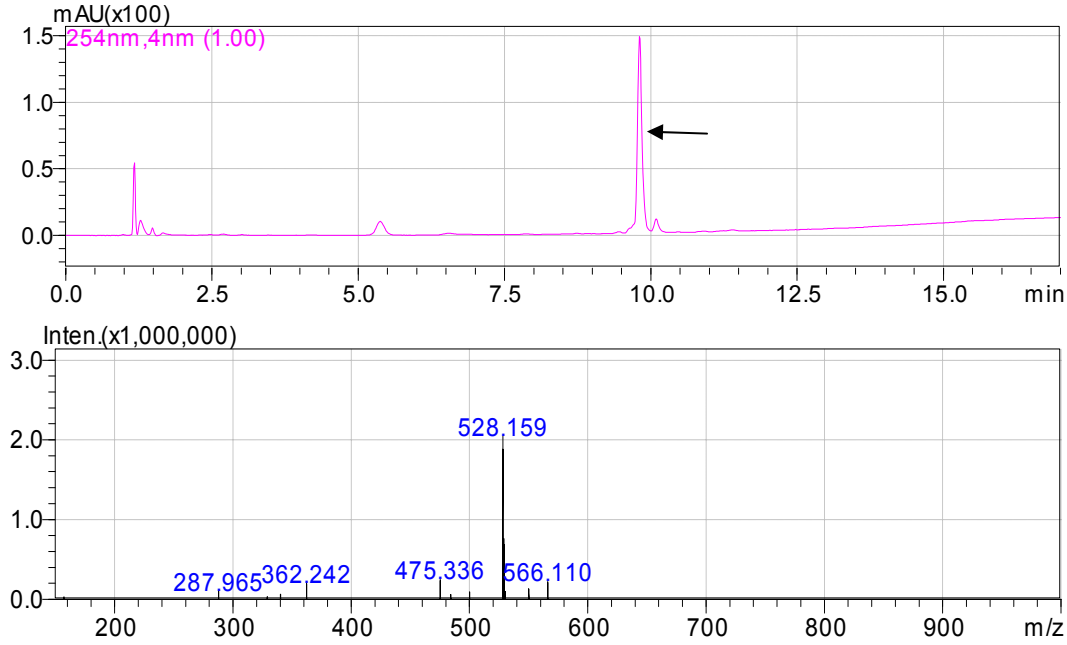


10.3 LC-MS spectra for selected compounds

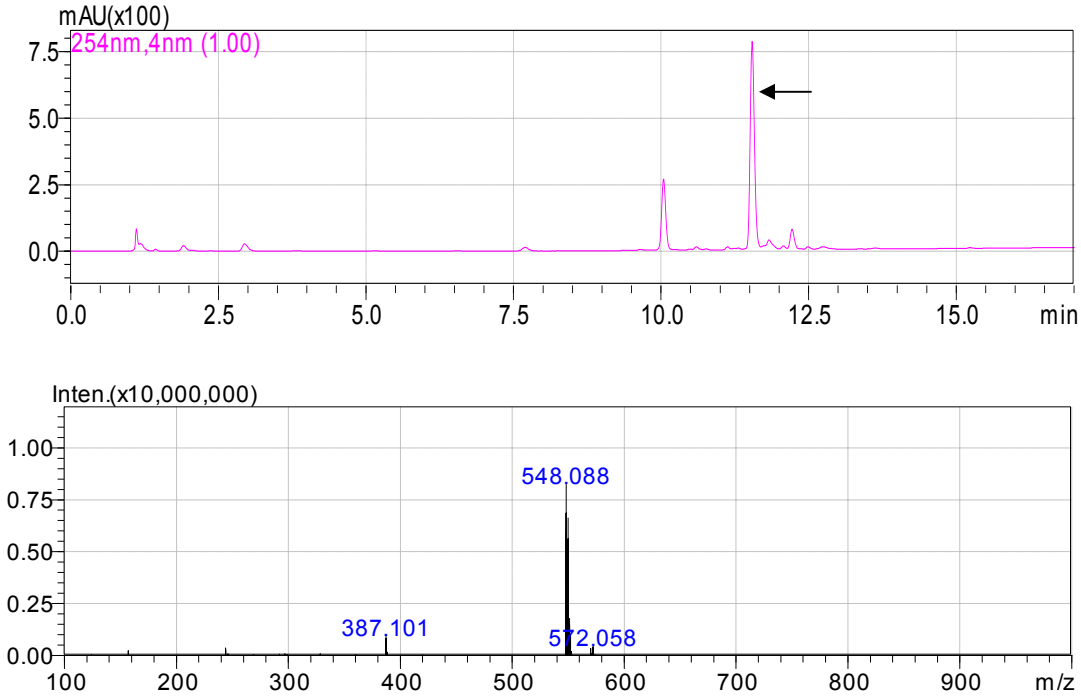
Compound 2-A-2-A3-2C



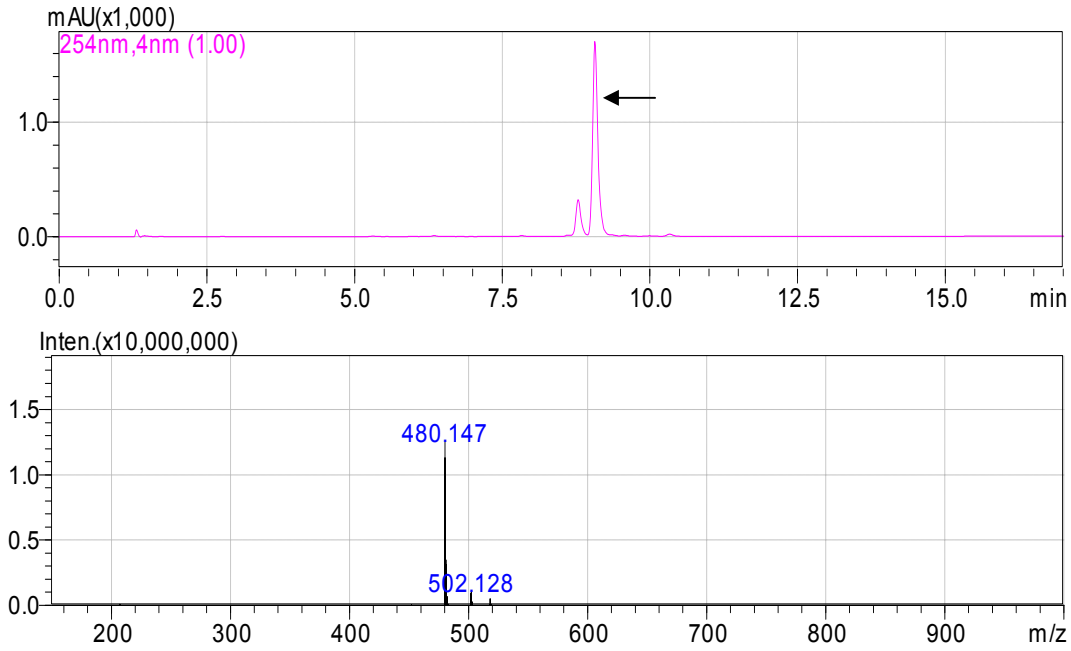
Compound 2-B-2-B5-2C



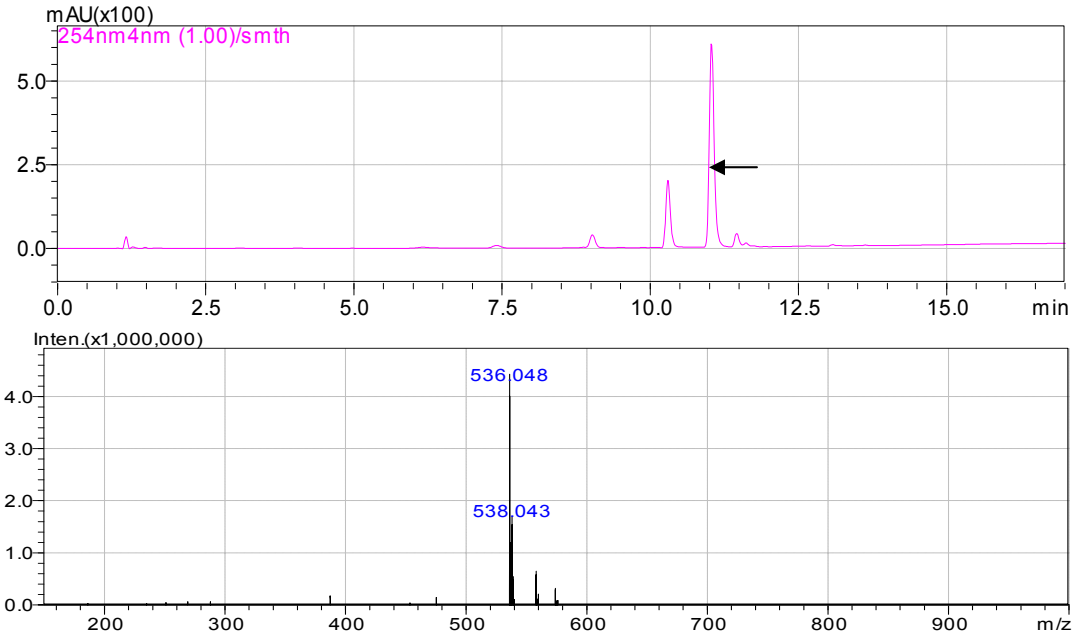
Compound 2-C-2-A10-5C



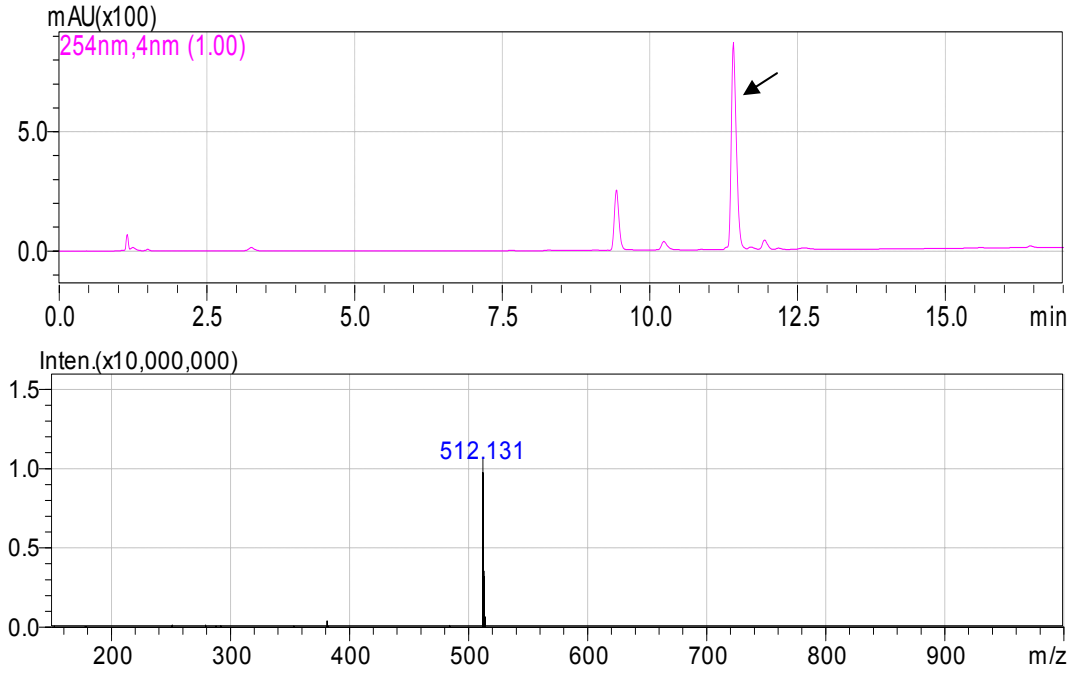
Compound 2-D-2-B4-2C



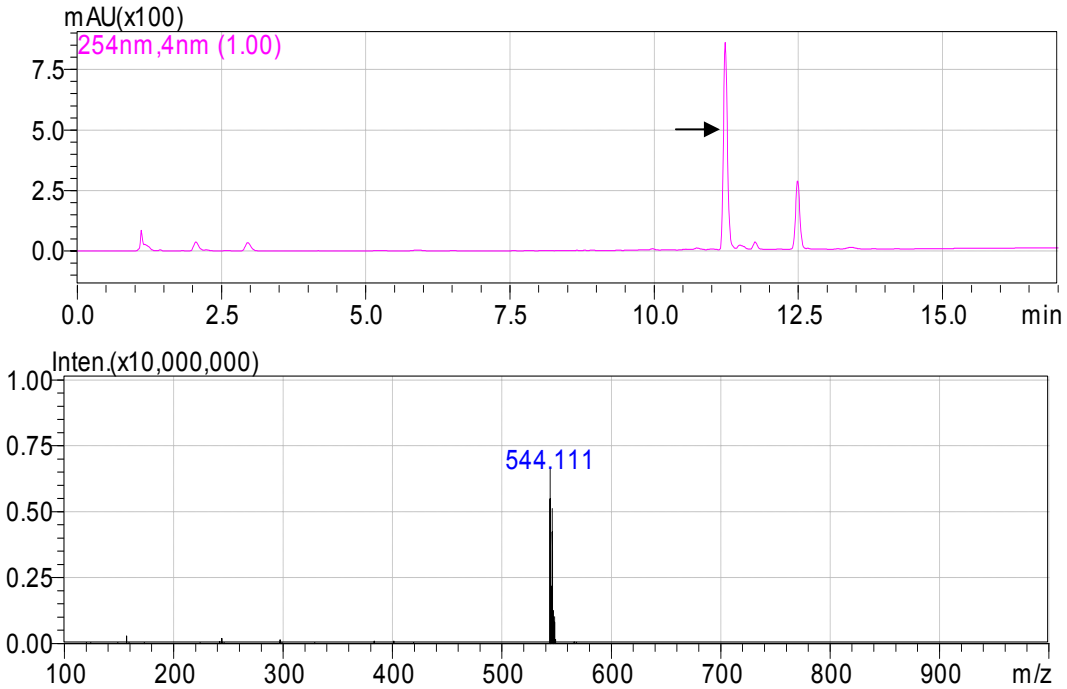
Compound 2-E-2-B1-2S



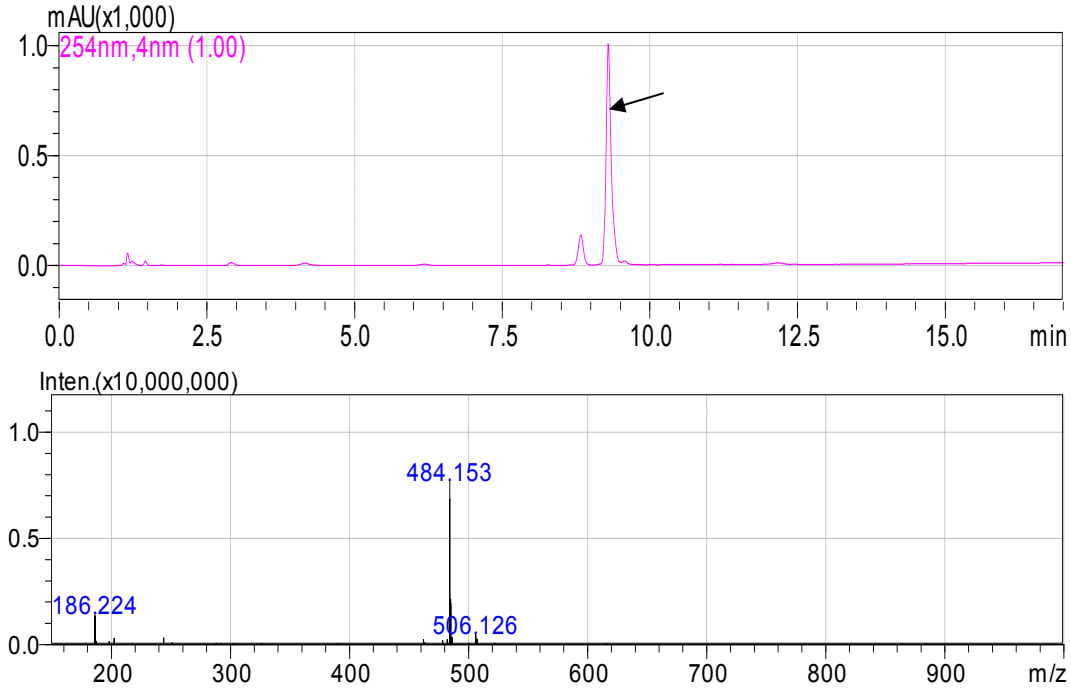
Compound 2-F-2-B1-2CAlk-6



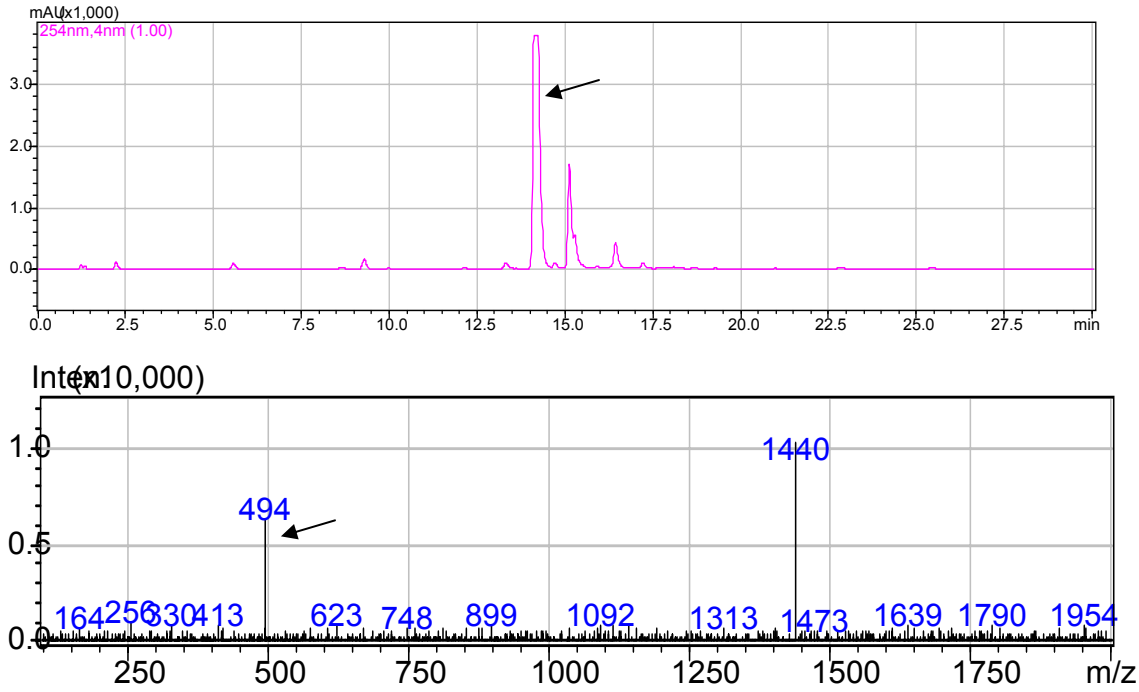
Compound 2-H-2-A7-5C



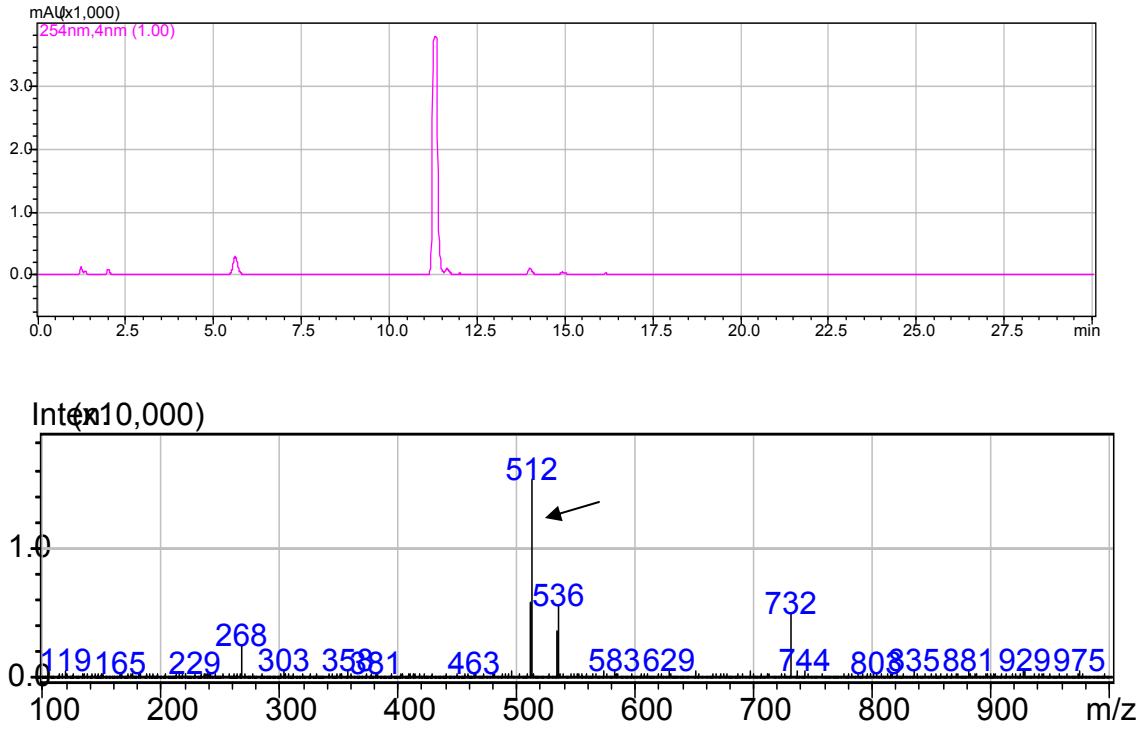
Compound 2-J-2-E4-4C



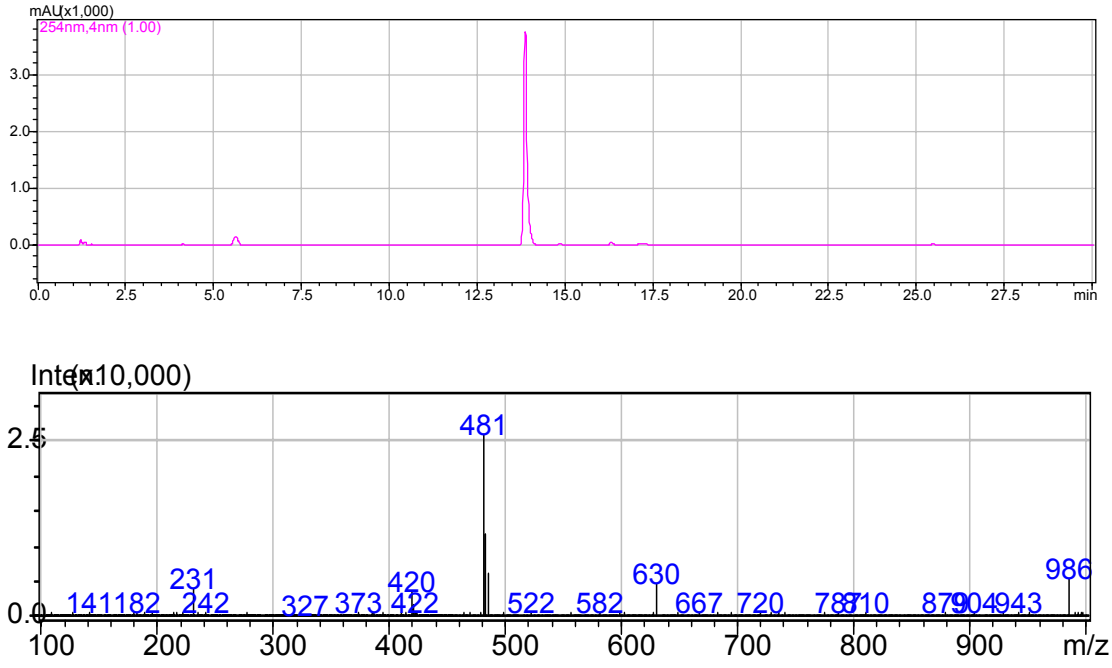
Compound 3-A26



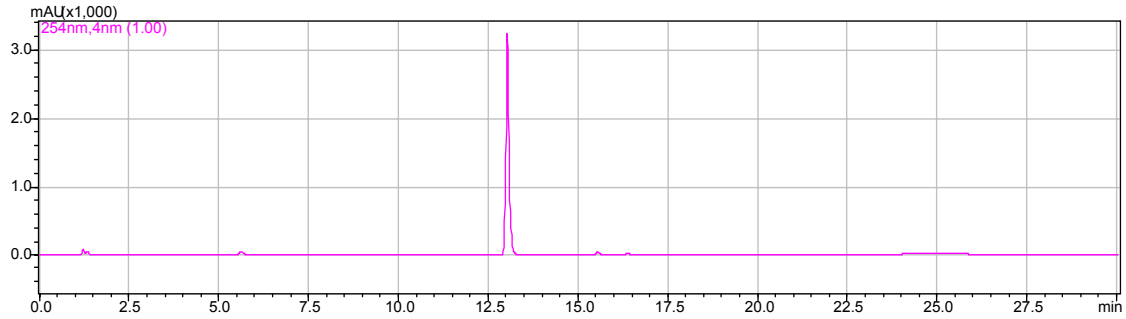
Compound 3-A33

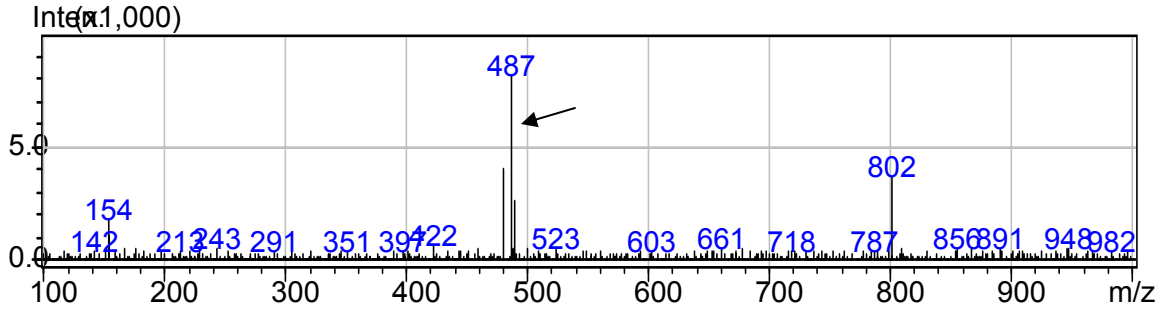


Compound 3-B15

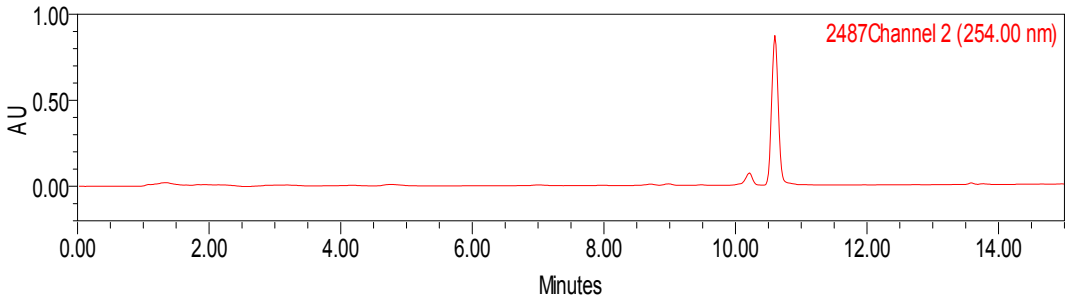


Compound 3-B23





Compound 4-1



Compound 4-6